



# A survey on recent trends in deep learning for nucleus segmentation from histopathology images

Anusua Basu<sup>1</sup> · Pradip Senapati<sup>1</sup> · Mainak Deb<sup>2</sup> · Rebika Rai<sup>3</sup> · Krishna Gopal Dhal<sup>1</sup>

Received: 3 February 2022 / Accepted: 13 February 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

## Abstract

Nucleus segmentation is an imperative step in the qualitative study of imaging datasets, considered as an intricate task in histopathology image analysis. Segmenting a nucleus is an important part of diagnosing, staging, and grading cancer, but overlapping regions make it hard to separate and tell apart independent nuclei. Deep Learning is swiftly paving its way in the arena of nucleus segmentation, attracting quite a few researchers with its numerous published research articles indicating its efficacy in the field. This paper presents a systematic survey on nucleus segmentation using deep learning in the last five years (2017–2021), highlighting various segmentation models (U-Net, SCPP-Net, Sharp U-Net, and LiverNet) and exploring their similarities, strengths, datasets utilized, and unfolding research areas.

**Keywords** Image segmentation · Nucleus segmentation · White blood cell segmentation · Histopathology image segmentation · Pathology image segmentation · Hematology image segmentation · Deep learning · Machine learning · Neural network · Deep neural network · Convolutional neural network · Cancer diagnosis

## 1 Introduction

Deep learning is a machine learning method that teaches computers to perform tasks that humans accomplish without thinking about them. A computer model can learn to carry out categorization tasks directly via images, text, or sound using deep learning. Modern precision can be attained by deep models, sometimes even outperforming human ability.

A sizable collection of labelled data and multi-layered neural network structures are used to train models. Cancer has historically been a fatal disease. It can be devastating even in today's technologically advanced world if it isn't caught in its earliest stages. Millions of lives could be saved by swiftly identifying any malignant cells. Nucleus segmentation is a method for identifying an image's nucleus by segmenting it into different parts. Deep learning is quickly gaining traction in the field of nucleus segmentation and has attracted quite a few researchers with its numerous published research articles demonstrating its usefulness.

Image Segmentation is principally a process that is used to partition a digital image into numerous segments or objects (Szeliski 2010). It is widely employed in several applications ranging from image compression (Rabani 2002) to medical image analysis (Ker et al. 2017) to robotic perception (Porzi et al. 2016). Image segmentation is categorized as semantic (Ahmed et al. 2020) and instance segmentation (Birodkar et al. 2021). Semantic segmentation groups together parts of an image that belong to the same class. Instance segmentation, which combines object detection and semantic segmentation, finds objects in well-defined categories. Medical image segmentation similar as natural image segmentation refers to the procedure of mining the anticipated object (organ) from a medical image that

---

✉ Krishna Gopal Dhal  
krishnagopal.dhal@midnaporecollege.ac.in

Anusua Basu  
mail2anusuacse@gmail.com

Pradip Senapati  
pradipsenapati2002@gmail.com

Mainak Deb  
mainakdeb15@gmail.com

Rebika Rai  
rrai@cus.ac.in

<sup>1</sup> Department of Computer Science and Application, Midnapore College (Autonomous), Paschim Medinipur, Midnapore, West Bengal, India

<sup>2</sup> Wipro Technologies, Pune, Maharashtra, India

<sup>3</sup> Department of Computer Applications, Sikkim University, Sikkim, India

can be instigated manually, semi-automatically or automatically intending to make anatomical or pathological structures transform indistinctly of the underlying images. Quite a few medical image segmentations take into account Breast and Breast Histopathology Segmentation (Liu et al. 2018a), liver and liver-tumour segmentation (Li 2015) (Vivanti et al. 2015), cell segmentation (Song et al. 2017) etc. as an input imagery and further applies mechanism into it. Medical image segmentation is a key part of Computer-Aided Diagnosis (CAD) and smart medicine, where features are taken from segmented images. Due to the rapid growth of deep learning techniques (Krizhevsky et al. 2017), medical image segmentation is no longer limited to hand-crafted features. Instead, Convolutional Neural Networks (CNN) can efficiently create hierarchical image features, which leads to the most accurate image segmentation models on popular benchmarks. This CNN method has inspired academics to develop deep learning segmentation models for histopathology images. This article focuses on recent trends in Deep Learning for Nucleus Segmentation from Histopathology Images throughout 2017–2021 by discussing U-Net (Ronneberger et al. 2015), SCPP-Net (Chanchal et al. 2021b), Sharp U-Net (Zunair and Hamza 2021), and LiverNet (Aatresh et al. 2021a) etc.

In recent years, Deep Learning-based innovative algorithms have shown state-of-the-art performance in medical imaging segmentation, processing, detection, and classification. The literature review was used to choose the four segmentation models. Only these four models were selected because they have demonstrated excellent nucleus segmentation performance in recent years. This introduction's references were chosen because they accurately represent.

The remaining sections of the paper are schematized as follows: Sect. 2 deliberates upon the importance of nucleus segmentation like cell counting, movement tracking, and morphological study, etc., stressing certain challenges while dealing with the same. The review and discussion on recent trends in Deep Learning for Nucleus Segmentation since 2017 is offered in Sect. 3. In Sect. 4, we also have an analysis based on year-wise published papers, backbone, loss functions for the research initiative since 2017 with a graphical representation based on the most frequently used backbone, loss function, optimizer, dataset etc. over the last five years in the literature survey. The architecture and brief description of some segmentation models (U-Net, SCPP-Net, Sharp U-Net, and LiverNet) along with their loss function and segmentation quality parameters have been conveyed in Sect. 5. Experimental datasets, training and implementation, and comparison of a few segmentation models along with the experimental outcomes are emphasised in Sect. 6 with some graphical representation based on segmentation results and

training loss. Lastly, Sect. 7 discusses the conclusion and future research directions.

## 2 Nucleus segmentation: need and challenges

This section briefly presents the need for and challenges of nucleus segmentation from histopathology images.

### 2.1 Need of nucleus segmentation

Segmenting cell nuclei in histopathology images is the preliminary step in analyzing current imaging data for biological and biomedical purposes. The fundamental steps for nucleus segmentation namely Cell counting (Grishagin 2015), Movement tracking (Dewan et al. 2011), Computational pathology (Louis et al. 2015), Cytometric analysis (Yu et al. 1989), Computer-Aided diagnosis (Kowal and Filipeczuk 2014) and Morphological study (Abdolhoseini et al. 2019) plays a dynamic role in analysing, diagnosing and grading cancerous cell. These fundamental steps we described as below:

- a. **Cell Counting:** It is a subclass of cytometry considered one of the methods used for counting or quantification of similar cells and is widely employed in numerous research and clinical practices. Superior quality microscopy images can be used with statistical classification algorithms for cell counting and recognition as part of image analysis (Han et al. 2012) performed off-line, keeping the error rate constant (Han et al. 2008).
- b. **Movement Tracking:** Automated tracking and analysis (Meijering et al. 2009) is seen as an important part of biomedical research, both for biological processes and for diagnosing diseases.
- c. **Computational Pathology:** It deals with analysing digitized pathology images with allied metadata wherein nucleus segmentation in digital microscopic tissue images aids high-quality features extraction for nucleus morpho metrics in it (Kumar et al. 2017).
- d. **Cytometric Analysis:** Nucleus segmentation is a significant step in the pipeline of many cytometric analyses. It has been used in a few studies to analyse the nucleus DNA to observe the association between the DNA ploidy pattern and the 5-year survival rate of advanced gastric cancer patients using paraffin-embedded tissue specimens (Kimura and Yonemura 1991).
- e. **Computer-Aided Diagnosis (CAD):** Computer-aided detection, also called CAD, is a useful tool for precise

diagnosis and prognosis (Su et al. 2015). It helps doctors interpret medical images.

- f. **Morphological Study:** This complex biological mechanism regulates cell proliferation, differentiation, development and disease (Jevtic et al. 2014). Cell morphology, for example, requires nucleus segmentation as a fundamental step because it provides valuable information about nucleus morphology, chromatin, DNA content, etc.

## 2.2 Challenges of nucleus segmentation

Dependent on a variability of measures like nuclides, malignant tumours, their life cycles etc., nuclei appear in different shapes and sizes. Several types of nuclei exist; however, lymphocyte nuclei (LN) are inflammatory nuclei having a regular shape, which have a major role in the immune system, and epithelial nuclei (EN) (Irshad et al. 2013) have nearly uniform chromatin distribution with a smooth boundary, which are the types of interest. Automated nuclei segmentation, though, is a well-researched problem in the field of digital pathology, but segmenting the nucleus turns out to be difficult due to the presence of a variety of blood cells. Furthermore, due to variability induced by elements in slide preparation (dyes concentration, damage of the given tissue sample, etc.) and image acquisition (digital noise existence, explicit features of the slide scanner, etc.), existing methods are unfitting and cannot be applied to all types of histopathology images (Hayakawa et al. 2021). Additionally, some of significant challenges that arise while segmenting nuclei are presented below:

1. There is a high level of heterogeneity in appearance between different types of organs or cells. So, methods that were made based on what was already known about geometric features can't be used right away on different images.
2. Nuclei are often clustered with many overlapping instances. Separating the clustered nuclei frequently necessitates additional processing.
3. In out-of-focus images, the boundary of nuclei seems blurry. That increases the difficulty of extricating dense illustrations from images. Furthermore, the factors that make the segmentation task difficult are the appearance of the nucleus and the noticeable variation in its shape.

An effective image processing approach must be able to overcome the aforesaid obstacles and challenges while maintaining the quality and accuracy of the underlying images in various situations.

## 3 Survey on deep learning based nucleus segmentation

For a few years, deep learning models have proven to be effective, vigorous, and accurate in biomedical image segmentation, specifically nucleus segmentation. This section includes a literature review of work done from 2017 to 2021 on Convolutional Neural Network (CNN) Model for nucleus segmentation, as shown in Table 1. The mentioned papers have been collected from the following sources:

- a. Google Scholar—<https://scholar.google.com>
- b. IEEE Xplore—<https://ieeexplore.ieee.org>
- c. ScienceDirect—<https://www.sciencedirect.com>
- d. SpringerLink—<https://www.springerlink.com>
- e. ACM Digital Library—<https://dl.acm.org>
- f. DBLP—<https://dblp.uni-trier.de>

Each of these above sources is queried with the following combinations of keywords:

- KW1: Deep Learning based histopathology Image Segmentation.  
 KW2: Deep Learning based hematology Image Segmentation.  
 KW3: Deep Learning based pathology Image Segmentation.  
 KW4: Deep learning based white blood cell segmentation.  
 KW5: Nucleus segmentation using deep learning.  
 KW6: Nucleus segmentation using machine learning.  
 KW7: Nucleus segmentation using Convolutional Neural Network.  
 KW8: White blood cell segmentation using Convolutional Neural Network.  
 KW9: Deep Neural Network based image segmentation.

## 4 Analysis and discussion

This section presents an analysis based on the reports on nucleus segmentation using CNN models which are reported in Table 1. The analysis is based on the year wise publications, datasets, CNN models, utilized segmentation metrics etc.

### 4.1 Analysis based on publication year

This subsection will present the analysis based on the publication years of the various works that were taken into consideration that are associated with nucleus segmentation.

**Table 1** Literature survey on deep learning models for nucleus segmentation during the year 2017 to 2021

Paper details	
<b>Year: 2017</b> Kumar et al. (2017) proposed a 3 layer CNN-3 for Generalized Nuclear Segmentation for Computational Pathology	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned The proposed model introduced deep learning-based segmentation technique for identifying nuclear boundaries (between touching and overlapping) for diverse datasets</p> <p><b>Comparison:</b> Cell Profiler (CP), Fiji and CNN Model with Two Convolutional Layer (CNN-2)</p> <p><b>Dataset:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were extracted from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Aggregated Jaccard Index (AJI), Average Hausdorff Distance, Average Dice's Coefficient (ADC) and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models in regard to diffused-chromatin and crowded nuclei of the breast, prostate, and colon. The measured values of AJI = 0.5083, Average Hausdorff Distance = 7.6615, ADC = 0.7623, and F1-Score = 0.8267</p> <p><b>Limitations:</b> In any scenario where the magnification is increased, more adjustments to the network architecture will be required in order to scan windows of greater size and take into account nuclei and the spatial context. If nuclear categorization is also wanted, it may be necessary to make some changes so that a variable number of pixel classes can be used</p>
<b>Year: 2018</b> Liu et al. (2018b) proposed a Mask Regional Convolutional Neural Network (MaskR-CNN) combined with Local Fully Connected Conditional Random Field (LFCCRF)	<p><b>Features:</b> <b>Backbone:</b> ResNet <b>Loss:</b> Classification Loss is Log loss and Regression Loss was Smooth L1 The proposed method employed a Feature Pyramid Network (FPN) based on the ResNet to gain stronger semantic features to localize the cervical nuclear boundary</p> <p><b>Comparison:</b> Multi-scale Watershed + Binary Classifier, Radiating Gradient Vector Flow (RGVF), Patch-based Fuzzy C-Means (Patch-based FCM) and Fully Convolutional Networks and Graph (FCN-G)</p> <p><b>Datasets:</b> 917 Pap-smear cells images extracted from Harlev dataset (Jantzen et al. 2005)</p> <p><b>Parameters:</b> Precision, Recall and Zydenbos Similarity Index (ZSI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby generating superior nuclear segmentation and further employing prior pixel level information for coarse segmentation. The measured values of Precision = 0.96 ± 0.05, Recall = 0.96 ± 0.11, ZSI = 0.95 ± 0.10</p> <p><b>Limitations:</b> The accuracy of this kind of wrong nuclear segmentation needs to be improved because it is clinically important</p>
<b>Year: 2018</b> Zhou et al. (2018) proposed a U-Net + + model for Medical Image Segmentation	<p><b>Features:</b> <b>Backbone:</b> U-Net <b>Loss:</b> combination of binary cross-entropy and dice coefficient The proposed model was a deeply-supervised encoder-decoder network with sub-networks connected by nested, solid pathways to reduce the semantic gap between feature maps</p> <p><b>Comparison:</b> U-Net and Wide U-Net</p> <p><b>Datasets:</b> Cell images of three distinct organ (colon, liver, lung) were extracted from Cell nuclei Dataset (Caicedo et al. 2019), Colon Polyp Dataset (Tajbakhsh et al. 2016; Zhou et al. 2017), Liver Dataset (Bernal et al. 2017) and Lung Nodule Dataset (Armato et al. 2011)</p> <p><b>Parameters:</b> Intersection over Union (IoU)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, highlighting that U-Net + + with deep supervision accomplished an average IoU improvement of 3.9 and 3.4 over U-Net, and wide U-Net respectively. IoU values (in %) obtained for cell nuclei, colon polyp, liver and lung nodule datasets were 92.63, 33.45, 82.90, and 77.21 respectively</p> <p><b>Limitations:</b> Deep supervision significantly improves the categorization of liver and lung tumors, but it has a bad impact on the segmentation of cell nuclei or colon polyps. This is due to the fact that the liver and polyps can be seen in different sizes in CT images and video frames</p>

Table 1 (continued)

Paper details	
<b>Year: 2019</b> Jung et al. (2019) proposed nuclei segmentation method based on Deep Convolutional Neural Networks (DCNNs) for histopathology images	<p><b>Backbone:</b> Not Mentioned</p> <p><b>Loss:</b> Log-likelihood loss function</p> <p>The Adam optimizer was employed to train the network. The proposed model used four major steps, namely pre-processing, colour normalization, nuclei segmentation, and post-processing, which were used in the proposed method. Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>) was used for nuclei segmentation</p> <p><b>Comparison:</b></p> <p>(i) <b>Evaluated on the MOSID dataset:</b> Cell Profiler (<b>CP</b>), Fiji, CNN Model with Two Convolutional Layers (<b>CNN-2</b>) and CNN Model with Three Convolutional Layer (<b>CNN-3</b>)</p> <p>(ii) <b>Evaluated on the BNS dataset:</b> PANGNET, Deconvolutional Network (<b>DeConv-Net</b>), Fully Convolutional Network (<b>FCN</b>), and Ensemble</p> <p><b>Datasets:</b> Cell images of three distinct organs (colon, prostate, liver) were extracted from Multiple organ H&amp;E-stained Histopathology Image Dataset (<b>MOSID</b>) (Kumar et al. 2017) and Breast Cancer Histopathology Image Dataset (<b>BNS</b>) (Naylor et al. 2017)</p> <p><b>Parameters:</b> Precision, Recall, F1 score, Aggregated Jaccard Index (<b>AJI</b>) and Average Dice's coefficient (<b>ADC</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models by employing Mask R-CNN with colour normalization and multiple inferences post-processing, providing robust nuclei segmentation results. The measured values were Precision = 0.913, Recall = 0.894, F1-Score = 0.861, ADC = 0.812, AJI = 0.669 for MOSID and Precision = 0.920, Recall = 0.923, F1-Score = 0.913, ADC = 0.838, AJI = 0.688 for the BNS Dataset, respectively</p> <p><b>Limitations:</b> The approach is mostly based on hand-made characteristics, which are limited by human thresholding, and the effectiveness of pixel-level segmentation cannot be fully evaluated using the ADC evaluation metrics</p> <p><b>Features:</b> <b>Backbone:</b> U-Shape Network <b>Loss:</b> not mentioned</p> <p>The proposed model employed blocks of the U-Sharped Network to transfer the information of the feature efficiently. The proposed model was prearranged into a multi-path fashion to train the network with different settings via a majority voting strategy to improve segmentation</p> <p><b>Comparison:</b> Unsupervised, Fuzzy C-means (<b>FCM</b>), Pairwise Markov's Random fields (<b>P-MRF</b>) and Shape Priors with Convolutional Neural network (<b>SP-CNN</b>)</p> <p><b>Datasets:</b> 917 images of Pap-smear cells extracted from Herlev dataset (Jantzen et al. 2005)</p> <p><b>Parameters:</b> Zydenbos Similarity Index (<b>ZSI</b>), Precision and Recall</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models with the usage of the D-MEM model on the Herlev dataset, thereby paving the way to be further extended to be applied for solving other variants of medical image segmentation. The measured values of ZSI = 0.933 ± 0.14, Precision = 0.946 ± 0.06 and Recall = 0.984 ± 0.00</p> <p><b>Limitations:</b> Since the method is based on a U-shaped network, a skip connection arises. The model would not have performed better due to the skip connection</p> <p><b>Features:</b> <b>Backbone:</b> ResNeXt-101-64 × 4d</p> <p><b>Loss:</b> Classification Loss was Cross-entropy loss and Regression Loss is Smooth L1</p> <p>The proposed model used a self-attention mechanism to capture the similarity between nuclei and to strongly handle moderately labelled training images, The framework wherein Region Proposal Network (<b>RPN</b>) to produce a large number of bounding box candidates to densely cover the image. Furthermore, an object score for an individual candidate is calculated</p> <p><b>Comparison:</b> DIST and CNN Model with Three Convolutional Layers (<b>CNN-3</b>)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were extracted from Nucleus Dataset (Cicconet et al. 2017), Haematoxylin and Eosin (H&amp;E)-stained Histopathology Dataset (Naylor et al. 2018)</p> <p><b>Parameters:</b> Aggregated Jaccard Index (<b>AJI</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models as it delivered a better solution when extreme segmentation accuracy was desirable by using it as a computer-assisted annotation tool to produce high-quality fully annotated training datasets. The measured value of AJI is 0.5353</p> <p><b>Limitations:</b> Incomplete annotations lower the classifier's performance because there are fewer training samples</p>
<b>Year: 2019</b> Zhao et al. (2019a) proposed a Deformable Multipath Ensemble Model ( <b>D-MEM</b> ) for automated segmentation of cervical nuclei in Pap smear Images	<p><b>Year: 2019</b> Feng et al. (2019) propose a <b>region-proposal module</b> to perform exemplar learning</p>



Table 1 (continued)

Paper details	
<b>Year: 2019</b> Allehabi et al. (2019) proposed approach combining Mask Regional Convolutional Neural Network (Mask R-CNN) and Visual Geometry Group-like Network (VGG-like Net) for Segmentation and Classification of Cervical Cells	<p><b>Features:</b> <b>Backbone:</b> ResNet-101 <b>Loss:</b> Not Mentioned The proposed model used RPN with Mask R-CNN with the purpose of generating image regions containing an object. Mask R-CNN was used for segmenting the whole cervical by dividing a cell into two areas, i.e., cytoplasm and nucleus, and the background. The VGG-like Net was utilised for cervical cell classification (2-class and 7-class classification problems)</p> <p><b>Comparison:</b> Multi-scale hierarchical Segmentation Algorithm, Mask Regional Convolutional Neural Network (Mask R-CNN)+Locally Fully Connected Conditional Random field (LFC-CRF), Radiating Gradient Vector Flow (RGVF), Fully Convolutional Networks and Graph (FCN-G), Fuzzy C-Means (FCM), Hard C-Means (HCM), Watershed and Neuromorphic graph Cut-based Segmentation (NGCS)</p> <p><b>Datasets:</b> Herlev Pap Smear Dataset (Jantzen et al. 2005), this dataset consists of 917 images of Pap-smear cells and Cell images of three distinct organs (Breast, Prostate and colon) were taken from Microsoft Common Objects in Context (MS COCO) Dataset (Lin et al. 2015)</p> <p><b>Parameters:</b> Precision, Recall and Zydenbos Similarity index (ZSI), Specificity</p> <p><b>Inference:</b> The proposed model was better in terms of performance however, but it required high processing speed when compared to other methods. The measured values of Precision = <math>0.92 \pm 0.06</math>, Recall = <math>0.91 \pm 0.05</math>, ZSI = <math>0.91 \pm 0.04</math> and Specificity = <math>0.83 \pm 0.10</math></p> <p><b>Limitations:</b> The work's drawback is the requirement for more processing power than the alternative solutions</p> <p><b>Features:</b> <b>Backbone:</b> U-Net <b>Loss:</b> Not Mentioned The proposed model employed two paradigms to extract image features at different scales, with 13 M total parameters. With the optimized progressive growing training method, the residual module was added to the extended path of the U-net structure. This helped solve problems in other fields with detecting and separating targets of different sizes</p> <p><b>Comparison:</b> Unsupervised, Fuzzy C-means (FCM), Shape Priors with Convolutional Neural network (SP-CNN) and Densely Connected U-Net (Dense U-Net)</p> <p><b>Datasets:</b> 917 images of Pap-smear cells extracted from Herlev dataset (Jantzen et al. 2005)</p> <p><b>Parameters:</b> Zydenbos Similarity Index (ZSI), Precision and Recall</p> <p><b>Inference:</b> The proposed model was better in terms of performance, also highlighting PGU-net+ had superior accuracy for cells of different sizes and shapes and is effective for extracting multi-scale information, making the task of extracting multi-scale information more explicit as compared to other models. The measured values of ZSI = <math>0.925 \pm 0.09</math>, Precision = <math>0.901 \pm 0.13</math>, Recall = <math>0.968 \pm 0.04</math></p> <p><b>Limitations:</b> The network is limited by the fact that it uses fixed-size receiving fields for things of different sizes</p> <p><b>Features:</b> <b>Backbone:</b> ResNet-50 <b>Loss:</b> Classification Loss was Binary Cross-entropy (BCE) loss and Regression Loss is Smooth L1 A Stochastic Gradient Descent (SGD) optimizer with a momentum of 0.9Nu was used. The proposed model used nuclei masks for training by evaluating objects and bounding boxes; region-based convolutional neural networks (R-CNN) to identify cell nuclei in oral histological tissues, and the network was pre-trained on the ImageNet dataset and fine-tuned using their dataset</p> <p><b>Comparison:</b> Expectation Maximization—Gaussian Mixture Model (EN-GMM), K-Means and Semantic Segmentation (SegNet)</p> <p><b>Datasets:</b> The proposed model used dataset built from tongue slides of 30 mice previously submitted to a carcinogen during two experiments performed between 2009 and 2010, duly approved by the Ethics Committee on the Use of Animals under protocol number 038/09 at the Federal University of Uberlandia, Brazil</p> <p><b>Parameters:</b> Accuracy (<math>A_{CC}</math>), Sensitivity (<math>S_p</math>), Specificity (<math>S_p</math>), Correspondence Rate (<math>C_R</math>) and Dice Coefficient (DC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models identified using qualitative and quantitative analysis. The measured values of <math>A_{CC} = 89.52 \pm 0.04</math>, <math>C_R = 0.76 \pm 0.10</math>, <math>D_c = 0.84 \pm 0.06</math></p> <p><b>Limitations:</b> The work hasn't yet looked into how to divide up images of oral dysplasia</p>
<b>Year: 2019</b> Zaho et al. (2019) proposed a Progressive Growing U-Net (PGU-Net+) for Automated Cervical Nuclei Segmentation	<p><b>Features:</b> <b>Backbone:</b> U-Net <b>Loss:</b> Not Mentioned The proposed model employed two paradigms to extract image features at different scales, with 13 M total parameters. With the optimized progressive growing training method, the residual module was added to the extended path of the U-net structure. This helped solve problems in other fields with detecting and separating targets of different sizes</p> <p><b>Comparison:</b> Unsupervised, Fuzzy C-means (FCM), Shape Priors with Convolutional Neural network (SP-CNN) and Densely Connected U-Net (Dense U-Net)</p> <p><b>Datasets:</b> 917 images of Pap-smear cells extracted from Herlev dataset (Jantzen et al. 2005)</p> <p><b>Parameters:</b> Zydenbos Similarity Index (ZSI), Precision and Recall</p> <p><b>Inference:</b> The proposed model was better in terms of performance, also highlighting PGU-net+ had superior accuracy for cells of different sizes and shapes and is effective for extracting multi-scale information, making the task of extracting multi-scale information more explicit as compared to other models. The measured values of ZSI = <math>0.925 \pm 0.09</math>, Precision = <math>0.901 \pm 0.13</math>, Recall = <math>0.968 \pm 0.04</math></p> <p><b>Limitations:</b> The network is limited by the fact that it uses fixed-size receiving fields for things of different sizes</p> <p><b>Features:</b> <b>Backbone:</b> ResNet-50 <b>Loss:</b> Classification Loss was Binary Cross-entropy (BCE) loss and Regression Loss is Smooth L1 A Stochastic Gradient Descent (SGD) optimizer with a momentum of 0.9Nu was used. The proposed model used nuclei masks for training by evaluating objects and bounding boxes; region-based convolutional neural networks (R-CNN) to identify cell nuclei in oral histological tissues, and the network was pre-trained on the ImageNet dataset and fine-tuned using their dataset</p> <p><b>Comparison:</b> Expectation Maximization—Gaussian Mixture Model (EN-GMM), K-Means and Semantic Segmentation (SegNet)</p> <p><b>Datasets:</b> The proposed model used dataset built from tongue slides of 30 mice previously submitted to a carcinogen during two experiments performed between 2009 and 2010, duly approved by the Ethics Committee on the Use of Animals under protocol number 038/09 at the Federal University of Uberlandia, Brazil</p> <p><b>Parameters:</b> Accuracy (<math>A_{CC}</math>), Sensitivity (<math>S_p</math>), Specificity (<math>S_p</math>), Correspondence Rate (<math>C_R</math>) and Dice Coefficient (DC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models identified using qualitative and quantitative analysis. The measured values of <math>A_{CC} = 89.52 \pm 0.04</math>, <math>C_R = 0.76 \pm 0.10</math>, <math>D_c = 0.84 \pm 0.06</math></p> <p><b>Limitations:</b> The work hasn't yet looked into how to divide up images of oral dysplasia</p>
<b>Year: 2019</b> Silva et al. (2019) proposed a segmentation method for histopathological images of oral dysplasia as based on an artificial neural network model and post-processing stage	<p><b>Features:</b> <b>Backbone:</b> U-Net <b>Loss:</b> Not Mentioned The proposed model employed two paradigms to extract image features at different scales, with 13 M total parameters. With the optimized progressive growing training method, the residual module was added to the extended path of the U-net structure. This helped solve problems in other fields with detecting and separating targets of different sizes</p> <p><b>Comparison:</b> Unsupervised, Fuzzy C-means (FCM), Shape Priors with Convolutional Neural network (SP-CNN) and Densely Connected U-Net (Dense U-Net)</p> <p><b>Datasets:</b> 917 images of Pap-smear cells extracted from Herlev dataset (Jantzen et al. 2005)</p> <p><b>Parameters:</b> Zydenbos Similarity Index (ZSI), Precision and Recall</p> <p><b>Inference:</b> The proposed model was better in terms of performance, also highlighting PGU-net+ had superior accuracy for cells of different sizes and shapes and is effective for extracting multi-scale information, making the task of extracting multi-scale information more explicit as compared to other models. The measured values of ZSI = <math>0.925 \pm 0.09</math>, Precision = <math>0.901 \pm 0.13</math>, Recall = <math>0.968 \pm 0.04</math></p> <p><b>Limitations:</b> The network is limited by the fact that it uses fixed-size receiving fields for things of different sizes</p> <p><b>Features:</b> <b>Backbone:</b> ResNet-50 <b>Loss:</b> Classification Loss was Binary Cross-entropy (BCE) loss and Regression Loss is Smooth L1 A Stochastic Gradient Descent (SGD) optimizer with a momentum of 0.9Nu was used. The proposed model used nuclei masks for training by evaluating objects and bounding boxes; region-based convolutional neural networks (R-CNN) to identify cell nuclei in oral histological tissues, and the network was pre-trained on the ImageNet dataset and fine-tuned using their dataset</p> <p><b>Comparison:</b> Expectation Maximization—Gaussian Mixture Model (EN-GMM), K-Means and Semantic Segmentation (SegNet)</p> <p><b>Datasets:</b> The proposed model used dataset built from tongue slides of 30 mice previously submitted to a carcinogen during two experiments performed between 2009 and 2010, duly approved by the Ethics Committee on the Use of Animals under protocol number 038/09 at the Federal University of Uberlandia, Brazil</p> <p><b>Parameters:</b> Accuracy (<math>A_{CC}</math>), Sensitivity (<math>S_p</math>), Specificity (<math>S_p</math>), Correspondence Rate (<math>C_R</math>) and Dice Coefficient (DC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models identified using qualitative and quantitative analysis. The measured values of <math>A_{CC} = 89.52 \pm 0.04</math>, <math>C_R = 0.76 \pm 0.10</math>, <math>D_c = 0.84 \pm 0.06</math></p> <p><b>Limitations:</b> The work hasn't yet looked into how to divide up images of oral dysplasia</p>

Table 1 (continued)

Paper details	
<b>Year: 2019</b> Yoo et al. (2019) proposed a weakly supervised nuclei segmentation method, which requires only point annotations for training	<p><b>Backbone:</b> ResNet-18</p> <p><b>Loss:</b> Binary Cross-entropy (BCE)</p> <p>The Adam optimizer with an initial learning rate of 0.001 was employed. The proposed model introduced an auxiliary network, called <b>Pseudo EdgeNet</b>, that directs the segmentation network to identify nuclei edges even without edge annotations</p> <p><b>Comparison:</b> Baseline and Dense Conditional Random Field (<b>DenseCRF</b>)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were extracted from Multi-Organ Nuclei Segmentation (MoNuSeg) (Kumar et al. 2017, 2020) and Triple Negative Breast Cancer (TNBC) (Naylor et al. 2018) dataset</p> <p><b>Parameters:</b> Intersection over Union (<b>IoU</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models as Pseudo EdgeNet without edge annotations identified edges that act as a strong constraint for weakly-supervised learning. However, given the same amount of data, the performance of weakly-supervised learning was bound to that of supervised learning. MoNuSeg and TNBC Dataset measured IoU values of 0.6136 and 0.6038, respectively</p> <p>This work can't be annotated for a small number of masks</p> <p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Mean Squared Error (MSE) and Binary Cross-entropy (BCE)</p> <p>Adam optimisation with an initial learning rate of <math>10^{-4}</math> was used, and after 25 epochs, it was reduced to <math>10^{-5}</math>. The proposed model introduced a Pre-activated Residual Network with 50 layers (<b>Preact-ResNet50</b>) that was applied for feature extraction. Further, nearest neighbour up-sampling via three distinct branches, namely the Nuclear Pixel (NP) branch, the HoVer branch, and the Nuclear Classification (NC) branch, was employed to instantly attain correct nuclear segmentation and classification</p> <p><b>Comparison:</b> Fully Convolutional Neural Network (<b>FCN-8</b>), Segmentation Network (<b>SegNet</b>), U-Net, Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>), Deep Convolutional Auto-encoder Network (<b>DCAN</b>), Micro-Net and DIST</p> <p><b>Datasets:</b> Cell images of four distinct organs (Breast, Liver, Kidney, and Prostate) were taken from Combined CPM (Vu et al. 2018), Triple Negative Breast Cancer (<b>TNBC</b>) (Naylor et al. 2018), CoNSEP (Lu et al. 2018)</p> <p><b>Parameters:</b> Dice, Aggregated Jaccard Index (<b>AJI</b>), Detection Quality (<b>DQ</b>), Segmentation Quality (<b>SQ</b>), Panoptic Quality (<b>PQ</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models because of the interpretable and reliable evaluation framework that could excellently enumerate performance and overcome the limitations. The measured values of (Dice = 0.801, AJI = 0.626, DQ = 0.774, SQ = 0.778, PQ = 0.606), (Dice = 0.749, AJI = 0.590, DQ = 0.743, SQ = 0.759, PQ = 0.578), and (Dice = 0.664, AJI = 0.404, DQ = 0.529, SQ = 0.764, PQ = 0.408) for the Combined CPM, TNBC, and CoNSEP datasets</p> <p><b>Limitations:</b> The ratio of unsuccessful detections may rise when comparing techniques across several datasets, particularly on samples with a high concentration of difficult-to-identify nuclei, which may have a detrimental effect on the AJI assessment</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Smooth Jaccard and Mean Squared Stochastic Gradient Descent (SGD) optimization was employed with the model incorporating a smaller number of parameters (21 M). SpaNet was used in the proposed model to perform pixel-wise segmentation and centroid detection maps of nuclei prediction. The spectral clustering method was applied to the output of SpaNet</p> <p><b>Comparison:</b> CNN Model with Three Convolutional Layer (<b>CNN-3</b>), DR, Deep Convolutional Auto-encoder Network (<b>DCAN</b>), Path Aggregation Network (<b>PA-Net</b>), Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>), Biodiversity and Ecosystem Services Network (<b>BES-Net</b>) and Contour-aware Informative Aggregation Network (<b>CIA-Net</b>)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (Kidney, Stomach, Liver, Bladder, Colorectal, Prostate and Breast) were taken from Multi-organ Dataset (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (<b>AJI</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, as SpaNet leads to correct instances of positional information by incorporating a lesser number of parameters, giving it a better chance to generalise on unseen data. The measured values of AJI are 62.39% and 63.40% of seen and unseen organs, respectively, and F1-Score is 82.81% and 84.51% of seen and unseen organs for the multi-organ dataset</p> <p><b>Limitations:</b> In the proposed network, after the Down Transitioning Block (DTB) and Up Transitioning Block (UTB) units, feature aggregation is not employed since it loses direct access to the positional information</p>
<b>Year: 2019</b> Graham et al. (2019) proposed a Horizontal and Vertical Distance Network ( <b>HoVer-Net</b> ) on multiple H&E histology images	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Mean Squared Error (MSE) and Binary Cross-entropy (BCE)</p> <p>Adam optimisation with an initial learning rate of <math>10^{-4}</math> was used, and after 25 epochs, it was reduced to <math>10^{-5}</math>. The proposed model introduced a Pre-activated Residual Network with 50 layers (<b>Preact-ResNet50</b>) that was applied for feature extraction. Further, nearest neighbour up-sampling via three distinct branches, namely the Nuclear Pixel (NP) branch, the HoVer branch, and the Nuclear Classification (NC) branch, was employed to instantly attain correct nuclear segmentation and classification</p> <p><b>Comparison:</b> Fully Convolutional Neural Network (<b>FCN-8</b>), Segmentation Network (<b>SegNet</b>), U-Net, Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>), Deep Convolutional Auto-encoder Network (<b>DCAN</b>), Micro-Net and DIST</p> <p><b>Datasets:</b> Cell images of four distinct organs (Breast, Liver, Kidney, and Prostate) were taken from Combined CPM (Vu et al. 2018), Triple Negative Breast Cancer (<b>TNBC</b>) (Naylor et al. 2018), CoNSEP (Lu et al. 2018)</p> <p><b>Parameters:</b> Dice, Aggregated Jaccard Index (<b>AJI</b>), Detection Quality (<b>DQ</b>), Segmentation Quality (<b>SQ</b>), Panoptic Quality (<b>PQ</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models because of the interpretable and reliable evaluation framework that could excellently enumerate performance and overcome the limitations. The measured values of (Dice = 0.801, AJI = 0.626, DQ = 0.774, SQ = 0.778, PQ = 0.606), (Dice = 0.749, AJI = 0.590, DQ = 0.743, SQ = 0.759, PQ = 0.578), and (Dice = 0.664, AJI = 0.404, DQ = 0.529, SQ = 0.764, PQ = 0.408) for the Combined CPM, TNBC, and CoNSEP datasets</p> <p><b>Limitations:</b> The ratio of unsuccessful detections may rise when comparing techniques across several datasets, particularly on samples with a high concentration of difficult-to-identify nuclei, which may have a detrimental effect on the AJI assessment</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Smooth Jaccard and Mean Squared Stochastic Gradient Descent (SGD) optimization was employed with the model incorporating a smaller number of parameters (21 M). SpaNet was used in the proposed model to perform pixel-wise segmentation and centroid detection maps of nuclei prediction. The spectral clustering method was applied to the output of SpaNet</p> <p><b>Comparison:</b> CNN Model with Three Convolutional Layer (<b>CNN-3</b>), DR, Deep Convolutional Auto-encoder Network (<b>DCAN</b>), Path Aggregation Network (<b>PA-Net</b>), Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>), Biodiversity and Ecosystem Services Network (<b>BES-Net</b>) and Contour-aware Informative Aggregation Network (<b>CIA-Net</b>)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (Kidney, Stomach, Liver, Bladder, Colorectal, Prostate and Breast) were taken from Multi-organ Dataset (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (<b>AJI</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, as SpaNet leads to correct instances of positional information by incorporating a lesser number of parameters, giving it a better chance to generalise on unseen data. The measured values of AJI are 62.39% and 63.40% of seen and unseen organs, respectively, and F1-Score is 82.81% and 84.51% of seen and unseen organs for the multi-organ dataset</p> <p><b>Limitations:</b> In the proposed network, after the Down Transitioning Block (DTB) and Up Transitioning Block (UTB) units, feature aggregation is not employed since it loses direct access to the positional information</p>

Table 1 (continued)

Paper details	
<b>Year: 2019</b> Zeng et al. (2019) proposed a Unet-based neural network model, Residual Inception Channel Attention-UNet (RIC-UNet) for Nuclei Segmentation in Histology Images	<p><b>Features:</b> <b>Backbone:</b> U-Net <b>Loss:</b> Focal Loss Adam's optimizer was employed. The proposed model used DC blocks for up-sampling and for accurate nuclei segmentation (RIC-UNet), and residual blocks, multi-scale, and channel attention mechanisms were applied</p> <p><b>Comparison:</b> Cell Profiler (CP), Fiji, CNN Model with Two Convolutional Layer (CNN-2), CNN Model with Three Convolutional Layer (CNN-3) and U-Net</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from The Cancer Genome Atlas (TCGA) dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Dice, F1-score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, as it was capable of extracting features from images of different resolutions. The measured values of AJI are 0.5635, Dice is 0.8008, and 0.8278 is for the TCGA Dataset</p> <p><b>Limitations:</b> Although the research has enhanced the outcomes for nucleus segmentation, there is still much potential for improvement, particularly in certain challenging circumstances when the nuclei and contours are not entirely obvious. Despite the fact that the RIC-UNet approach has a higher discriminating impact than U-net on some deeper backgrounds with colours that are barely different from the nuclei's colors, some results are incorrectly segmented</p> <p><b>Features:</b> <b>Backbone:</b> Not-Mentioned <b>Loss:</b> Binary Cross-entropy (BCE) Adam's optimizer was employed to update the weights. To separate the nucleus from the background, the proposed model used U-Net to perform semantic segmentation. Regression U-Net was then applied to individual nuclei to generate its own distance map. Based on it, the watershed algorithm was used to generate the final segmentation mask. The number of trainable parameters for the first and second stages of the algorithm were identical at 1,941,105 for each stage</p> <p><b>Comparison:</b> U-Net, CNN Model with Three Convolutional Layer (CNN-3) and DR</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Average Dice Score, F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of overall performance (overall AJI); however, in terms of average Dice score and F1-score, the compared algorithms produced roughly equivalent results. The average dice score was measured to be 79.32, the F1-score was 81.88, and the AJI was 56.87</p> <p><b>Limitations:</b> The proposed model cannot be properly segmented for complex shapes and structural images</p> <p><b>Features:</b> <b>Backbone:</b> DenseNet <b>Loss:</b> Smooth Truncated and Soft Dice loss The Adam optimizer was used across the entire network, with the learning rate set to 0.001. The proposed module applied pyramidal features hierarchically by constructing multi-level lateral connections amongst encoders and decoders and further using a pyramidal feature extraction approach over the encoder structure</p> <p><b>Comparison:</b> Cell Profiler (CP), Fiji, CNN Model with Three Convolutional Layer (CNN-3), Deep Convolutional Auto-encoder Network (DCAN), Path Aggregation Network (PA-Net) and Biodiversity and Ecosystem Services Network (BES-Net)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from MoNuSeg Dataset (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby making the proposed model superior enough to be applied as a generalized model over wide varieties of medical image segmentation tasks as a performance booster. The measured values of AJI are 0.6129 and 0.6306 for the seen and unseen organs, and F1-Score is 0.8279 and 0.8458 for the seen and unseen organs in the MoNuSeg Dataset</p> <p><b>Limitations:</b> The suggested technique should not be well suited to a variety of situations, notably the diffuse chromatin and connected nuclei in hidden organs</p>
<b>Year: 2019</b> Mahbod et al. (2019) proposed a Two-Stage U-Net Algorithm for Segmentation of Nuclei in H&E-Stained Tissues images	<p><b>Features:</b> <b>Backbone:</b> Not-Mentioned <b>Loss:</b> Binary Cross-entropy (BCE) Adam's optimizer was employed to update the weights. To separate the nucleus from the background, the proposed model used U-Net to perform semantic segmentation. Regression U-Net was then applied to individual nuclei to generate its own distance map. Based on it, the watershed algorithm was used to generate the final segmentation mask. The number of trainable parameters for the first and second stages of the algorithm were identical at 1,941,105 for each stage</p> <p><b>Comparison:</b> U-Net, CNN Model with Three Convolutional Layer (CNN-3) and DR</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Average Dice Score, F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of overall performance (overall AJI); however, in terms of average Dice score and F1-score, the compared algorithms produced roughly equivalent results. The average dice score was measured to be 79.32, the F1-score was 81.88, and the AJI was 56.87</p> <p><b>Limitations:</b> The proposed model cannot be properly segmented for complex shapes and structural images</p> <p><b>Features:</b> <b>Backbone:</b> DenseNet <b>Loss:</b> Smooth Truncated and Soft Dice loss The Adam optimizer was used across the entire network, with the learning rate set to 0.001. The proposed module applied pyramidal features hierarchically by constructing multi-level lateral connections amongst encoders and decoders and further using a pyramidal feature extraction approach over the encoder structure</p> <p><b>Comparison:</b> Cell Profiler (CP), Fiji, CNN Model with Three Convolutional Layer (CNN-3), Deep Convolutional Auto-encoder Network (DCAN), Path Aggregation Network (PA-Net) and Biodiversity and Ecosystem Services Network (BES-Net)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from MoNuSeg Dataset (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby making the proposed model superior enough to be applied as a generalized model over wide varieties of medical image segmentation tasks as a performance booster. The measured values of AJI are 0.6129 and 0.6306 for the seen and unseen organs, and F1-Score is 0.8279 and 0.8458 for the seen and unseen organs in the MoNuSeg Dataset</p> <p><b>Limitations:</b> The suggested technique should not be well suited to a variety of situations, notably the diffuse chromatin and connected nuclei in hidden organs</p>
<b>Year: 2019</b> Zhou et al. (2019a) proposed a Contour-aware Informative Aggregation Network (CIA-Net) with multilevel information aggregation module between two task specific decoders and a novel smooth truncated loss	<p><b>Features:</b> <b>Backbone:</b> U-Net <b>Loss:</b> Focal Loss Adam's optimizer was employed. The proposed model used DC blocks for up-sampling and for accurate nuclei segmentation (RIC-UNet), and residual blocks, multi-scale, and channel attention mechanisms were applied</p> <p><b>Comparison:</b> Cell Profiler (CP), Fiji, CNN Model with Two Convolutional Layer (CNN-2), CNN Model with Three Convolutional Layer (CNN-3) and U-Net</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from The Cancer Genome Atlas (TCGA) dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Dice, F1-score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, as it was capable of extracting features from images of different resolutions. The measured values of AJI are 0.5635, Dice is 0.8008, and 0.8278 is for the TCGA Dataset</p> <p><b>Limitations:</b> Although the research has enhanced the outcomes for nucleus segmentation, there is still much potential for improvement, particularly in certain challenging circumstances when the nuclei and contours are not entirely obvious. Despite the fact that the RIC-UNet approach has a higher discriminating impact than U-net on some deeper backgrounds with colours that are barely different from the nuclei's colors, some results are incorrectly segmented</p> <p><b>Features:</b> <b>Backbone:</b> Not-Mentioned <b>Loss:</b> Binary Cross-entropy (BCE) Adam's optimizer was employed to update the weights. To separate the nucleus from the background, the proposed model used U-Net to perform semantic segmentation. Regression U-Net was then applied to individual nuclei to generate its own distance map. Based on it, the watershed algorithm was used to generate the final segmentation mask. The number of trainable parameters for the first and second stages of the algorithm were identical at 1,941,105 for each stage</p> <p><b>Comparison:</b> U-Net, CNN Model with Three Convolutional Layer (CNN-3) and DR</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Average Dice Score, F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of overall performance (overall AJI); however, in terms of average Dice score and F1-score, the compared algorithms produced roughly equivalent results. The average dice score was measured to be 79.32, the F1-score was 81.88, and the AJI was 56.87</p> <p><b>Limitations:</b> The proposed model cannot be properly segmented for complex shapes and structural images</p> <p><b>Features:</b> <b>Backbone:</b> DenseNet <b>Loss:</b> Smooth Truncated and Soft Dice loss The Adam optimizer was used across the entire network, with the learning rate set to 0.001. The proposed module applied pyramidal features hierarchically by constructing multi-level lateral connections amongst encoders and decoders and further using a pyramidal feature extraction approach over the encoder structure</p> <p><b>Comparison:</b> Cell Profiler (CP), Fiji, CNN Model with Three Convolutional Layer (CNN-3), Deep Convolutional Auto-encoder Network (DCAN), Path Aggregation Network (PA-Net) and Biodiversity and Ecosystem Services Network (BES-Net)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from MoNuSeg Dataset (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby making the proposed model superior enough to be applied as a generalized model over wide varieties of medical image segmentation tasks as a performance booster. The measured values of AJI are 0.6129 and 0.6306 for the seen and unseen organs, and F1-Score is 0.8279 and 0.8458 for the seen and unseen organs in the MoNuSeg Dataset</p> <p><b>Limitations:</b> The suggested technique should not be well suited to a variety of situations, notably the diffuse chromatin and connected nuclei in hidden organs</p>



Table 1 (continued)

Paper details	
<b>Year: 2019</b> Kang et al. (2019) proposed a novel nuclei segmentation approach based on a 2-stage learning framework and Deep Layer Aggregation (DLA) for nuclei segmentation in histopathological images	<p><b>Features:</b> <b>Backbone:</b> Not mentioned The ADDELTA optimizer was used, and U-Nets with DLA were extended by iteratively merging features across different levels. The proposed model used a two-step task by adding nuclei-boundary prediction (3 classes) as an intermediate step to convert the original binary segmentation, wherein estimation of nuclei and their boundaries was done in step 1, followed by generating a finely tuned segmentation map in step 2</p> <p><b>Comparison:</b> (i) <b>Evaluated on the TCGA dataset</b> – Fully Convolutional Network Layer-8 (FCN-8), Mask Regional Convolutional Neural Network (Mask R-CNN), U-Net, CNN Model with Three Convolutional Layer (CNN-3), DIST and Stacked U-Net (ii) <b>Evaluated on the TNBC dataset</b> – Deconvolutional Network (DeConv-Net), Fully Convolutional Network Layer-8 (FCN-8), Ensemble, U-Net and Stacked U-Net</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, Stomach) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomeczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016) and Triple-Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby making it applicable as a generalized model that could be applied across several types of cells, above all in various organs. The measured values of AJI=0.5895 and F1-Score=0.8079 for the TCGA Dataset and Recall=0.833, Precision=0.826, F1-Score=0.829, and AJI=0.611 for the TNBC Dataset, respectively</p> <p><b>Limitations:</b> According to the experiments, there isn't much of a performance difference between shallow and deep designs for the second stage, so the shallower one is selected when considering computing cost and efficiency</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Binary Cross-entropy (BCE) and Smooth L1 loss Stochastic Gradient Descent (SGD) with 0.9 momentum was applied as the optimizer. The proposed model used the Instance Relation Module (IRM) to formulate the cell association matrix for shifting information among discrete cell-instance features. Meanwhile, a sparsity-constrained Duplicate Removal Module (DRM) was anticipated to eradicate the misalignment while selecting the candidate among classification and localization accuracy</p> <p><b>Comparison:</b> Joint Optimization of Multiple Level Set (JOMLS), Cell Segmentation Proposal Network (CSP-Net) and Mask Regional Convolutional Neural Network (Mask R-CNN)</p> <p><b>Datasets:</b> Cervical Pap Smear (CPS) Dataset (Pissiti et al. 2018) with more than 8000 cell annotations in Pap smear image</p> <p><b>Parameters:</b> Average Jaccard Index (AJI) and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby indicating the proposed method could be applied as a generalized procedure and how reconnoitering instance relations could prove fruitful as well as influence its information in terms of feature representation, providing better semantic consistency. The measured values of (AJI =0.7185 and 0.5496) and (F1-Score=0.7497 and 0.7554) for cytoplasm and nuclei compare with the CPS Dataset</p> <p><b>Limitations:</b> In the dataset that was used, there were white blood cells and other complex background information. This meant that the algorithm had to be more resistant to noise</p>
<b>Year: 2019</b> Zhou et al. (2019b) proposed a novel Instance Relation Network (IR-Net) for robust overlapping Cervical cell segmentation	

Table 1 (continued)

Paper details	
<b>Year: 2019</b> Li et al. (2019a) proposed a bottom-up method for nuclear segmentation	<p><b>Features:</b> <b>Backbone:</b> Not clearly mentioned <b>Loss:</b> Cross Entropy (CE), Intersection Over Union (IOU) and Mean Square (MS) Stochastic Gradient Descent (SGD) optimization was used, and multiple thresholds were applied for the purpose of controlling the region growth algorithm. The proposed model used a Fully Convolutional Neural Network (FCN) for the sole purpose of semantic segmentation and for predicting Inside Mask, Center Mask, and Center Vector, which were useful for identifying Instance Mask</p> <p><b>Comparison:</b> CNN Model with Two Convolutional Layer (CNN-2) and CNN Model with Three Convolutional Layers (CNN-3)</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, and Stomach) were taken from dataset released by The Cancer Genome Atlas (TCGA) (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Dice and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance; however, in the scenario where different annotators bring to some extent different annotations, the sensitivity may make the model slightly complicated and difficult to understand and analyze. The measured values of AJI are 0.561 and Dice are 0.793 for the Kumar Dataset</p> <p><b>Limitations:</b> The centre vector is smooth inside a nuclear instance, which helps the model learn, but it changes quickly outside of nuclear instances, especially at the edges of nuclei that touch, which makes the model pay more attention to the edges</p>
<b>Year: 2018</b> Basha et al. (2018) proposed a Routine Colon Cancer Nuclei Network (RCC-Net) for histological routine colon cancer nuclei classification	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Categorical Cross-entropy Adam's optimizer was employed. The proposed model consisted of seven trainable layers with 1,512,868 learnable parameters, which outperformed Softmax CNN IN27 for the histological routine colon cancer nuclei classification task</p> <p><b>Comparison:</b> SoftmaxCNN_IN27, Softmax CNN, Alex-Net, CIFAR-VGG, GoogLeNet and Wide Residual Network (WRN)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Colon) were taken from CRC Histo-Phenotypes Dataset (Sirinukunwattana et al. 2016)</p> <p><b>Parameters:</b> Classification Accuracy and Weighted Average F1 Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance (test accuracy and weighted average F1 score) as compared to other models, thereby proving it efficient and generalized with regard to training time and data over fitting. The measured values of Classification Accuracy (training accuracy = 89.79% and testing accuracy = 80.61%) and Weighted Average F1-Score (training F1-Score = 0.9037 and testing F1-Score = 0.7887) for the CRCHisto Phenotypes Dataset</p> <p><b>Limitations:</b> The model is complex enough to give good results for standard histopathological images of colorectal cancer</p>
<b>Year: 2019</b> Wang et al. (2019a) proposed a Multi-Path Dilated Residual Network for Nuclei Segmentation and Detection	<p><b>Features:</b> <b>Backbone:</b> Dilated Residual Network (D-ResNet64) <b>Loss:</b> Binary Logarithmic Loss and Smooth L1 Loss Adam optimizer was employed. The proposed Mask Regional Convolutional Neural Network (Mask R-CNN) model was applied to the entire structure for the purpose of segmenting and detecting dense yet small objects to resolve one of the major issues of deep learning, which was information loss due to small objects</p> <p><b>Comparison:</b> Support Vector Machine (SVM), Random Forest, Logistic Regression (LR), UNet++ Morphology Post-processing, ResNet-50+Mask RCNN, ResNet101+Mask R-CNN, DenseNet121+Mask RCNN, ResNet50+Mask SSD and UNet+Deep Watershed Transform</p> <p><b>Datasets:</b> Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, Stomach) were taken from Data Science Bowl 2018 (DSB2018) Dataset (Caicedo et al. 2019; Data science bowl 2018), Multi-Organ Nuclei Segmentation (MoNuSeg) (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby making it a better model in terms of recognition and segmentation, especially for dense yet small objects, which was the major target of this paper. The measured values of AJI = 0.6145 and (AJI = 0.5128 and F1-Score = 0.7991) for the DSB 2018 and MoNuSeg datasets, respectively</p> <p><b>Limitations:</b> Pooling and other down-sampling operating methods can extend the perception in a dilated residual network, however, the feature map must be squeezed and the information deformation is lossy, thus the main image's structure information is lost</p>

Table 1 (continued)

Paper details	
<b>Year: 2020</b>	<b>Backbone:</b> Not clearly mentioned
Chen et al. (2020) proposed a Boundary-assisted Region Proposal Network ( <b>BRP-Net</b> ) that achieves robust instance-level nucleus segmentation	<b>Loss:</b> Focal Loss AdamW optimizers were employed for training. The proposed model introduced a task-aware feature encoding (TAFF) network, which was applied for semantic segmentation and detecting instance boundaries where the essential features were required to be extracted. Two stages of BRP-Net were considered: the first stage to acquire the instance proposal and the other for segmenting proposal-wise
<b>Comparison:</b>	(i) <b>Evaluated on the Kumar Dataset</b> – CNN Model with Three Convolutional Layer ( <b>CNN-3</b> ), DIST, Mask Regional Convolutional Neural Network ( <b>Mask R-CNN</b> ), Contour-aware Informative Aggregation Network ( <b>CIA-Net</b> ), Horizontal and Vertical Distance Network ( <b>HoVer-Net</b> ) and Spatial Pyramid Attention Network ( <b>Spa-Net</b> ) (ii) <b>Evaluated on the CPM17 Dataset</b> – Deep Residual Aggregation Network ( <b>DRAN</b> ), Horizontal and Vertical Distance Network ( <b>HoVer-Net</b> ) and Micro-Net
<b>Datasets:</b>	Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, Stomach) were taken from Multi-organ Nucleus Dataset ( <b>Kumar Dataset</b> ) (Kumar et al. 2017, 2020) and Computational Precision Medicine Dataset ( <b>CPM17</b> ) (Vu et al. 2018)
<b>Parameters:</b>	Aggregated Jaccard Index ( <b>AJI</b> ), F1-Score, Dice 1 and Dice 2
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, indicating that BRP-Net was robust to the variation of post-processing hyper-parameters and highly robust to the value of dilation radius. The measured values of AJI (%) = 64.22 and F1-Score (%) = 84.23 for the Kumar Dataset. And Dice 1 (%) has a measured value of 87.7. Dice 2 (%) = 79.5; F1-Score (%) = 73.1 for the CPM17 dataset
<b>Limitations:</b>	The disadvantage is the small number of open-source datasets used for training and testing
<b>Features:</b>	<b>Backbone:</b> Not clearly mentioned <b>Loss:</b> Cross-entropy Loss and Focal loss SGD minimised the loss function with momentum 0.9 and a batch size of 4. SUNets merge four parallel backbone nets using attention generation mechanisms. Stacked U-Net predicts pixel-wise nucleus segmentation, and stage 2 of SUNets takes the RGB value of the original image and outputs a binary map as input
<b>Comparison:</b>	Fully Convolutional Network Layer-8 ( <b>FCN-8</b> ), Mask Regional Convolutional Neural Network ( <b>Mask R-CNN</b> ), U-Net, CNN Model with Three Convolutional Layer ( <b>CNN-3</b> ), DIST, Stacked U-Net, U-Net (Deep Layer Aggregation or <b>DLA</b> ), Two-stage U-Net and Two-stage Learning U-Net ( <b>DLA</b> )
<b>Datasets:</b>	Cell images of seven distinct organs (breast, kidney, Liver, Prostate, Bladder, Colon, Stomach) were taken from The Cancer Genome Atlas ( <b>TCGA</b> ) Dataset (Tomczak et al. 2015); The Cancer Genome Atlas ( <b>TCGA</b> ) 2016) and Triple-Negative Breast Cancer ( <b>TNBC</b> ) Dataset (Naylor et al. 2018)
<b>Parameters:</b>	Precision, Recall, Aggregated Jaccard Index ( <b>AJI</b> ) and F1-Score
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thereby making the proposed model superior not just for the nuclei instances but also the cases wherein overlapped regions prevail. The measured values of AJI and F1-Score for the TCGA Dataset are 0.5965 and 0.8247, respectively. The measured values of precision = 0.853 and recall = 0.792. F1-Score = 0.806 and AJI = 0.621 for the CPM17 Dataset
<b>Limitations:</b>	Due to the skip connection in the model architecture, the proposed model does not perform well
<b>Features:</b>	<b>Backbone:</b> ResNet-101 <b>Loss:</b> Not mentioned Stochastic Gradient Descent (SGD) was used as an optimizer with an initial learning rate of 1e-1, a momentum of 0.99. The number of parameters in <b>PSPSegNet</b> exceeded 122 million. The proposed model employed data augmentation techniques, a training segmentation model, and post-processing steps to basically lessen over-fitting and boost generalization. The sparse stain color normalization method was used in the pre-processing step to lessen the color inconsistency between multi-institutional and multi-organ WSI images
<b>Comparison:</b>	Fully Convolutional Network ( <b>FCN</b> ), Fully Convolutional-DenseNet ( <b>FC DenseNet</b> ) and U-Net
<b>Datasets:</b>	Cell images of Four distinct organs (breast, kidney, Prostate, Stomach) were taken from Multi-Organ Nuclei Segmentation ( <b>MoNuSeg</b> ) (Kumar et al. 2017, 2020)
<b>Parameters:</b>	F1-Score and Aggregated Jaccard Index ( <b>AJI</b> )
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thereby indicating that the proposed PSPSegNet model could be effectively employed for the purpose of cell counting. The measured values of F1-Score and AJI are 0.8815 and 0.7080, respectively
<b>Limitations:</b>	Concerning the object-level AJI score, it should be noted that the model relies on fictitious training data, which could not yield extremely precise cell shapes
<b>Year: 2021</b>	<b>Backbone:</b> ResNet-101
Hassan et al. (2021a) proposed deep semantic nuclei segmentation model for multi-institutional WSI images and nuclei segmentation model (called Pyramid Scene Parsing with SegNet ( <b>PSPSegNet</b> ) for Multi-Institutional Histopathology Images	<b>Loss:</b> Not mentioned Stochastic Gradient Descent (SGD) was used as an optimizer with an initial learning rate of 1e-1, a momentum of 0.99. The number of parameters in <b>PSPSegNet</b> exceeded 122 million. The proposed model employed data augmentation techniques, a training segmentation model, and post-processing steps to basically lessen over-fitting and boost generalization. The sparse stain color normalization method was used in the pre-processing step to lessen the color inconsistency between multi-institutional and multi-organ WSI images

Table 1 (continued)

Paper details	
<b>Year: 2021</b> Lal et al. (2021) proposed a <b>NucleiSegNet</b> for Nuclei Segmentation of Liver Cancer Histopathology Images	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Dice and Jaccard loss An Adam optimizer was employed to calculate the optimal weights during back propagation. The total number of parameters was 15.7 million. The proposed architecture comprised three basic building blocks: a robust residual block basically for high-level semantic map extraction, a bottleneck block, and an attention decoder block for efficient object localization by reducing false positives</p> <p><b>Comparison:</b> CNN Model with One Convolutional Layer (<b>CNN-1</b>), CNN Model with Two Convolutional Layer (<b>CNN-2</b>), CNN Model with Three Convolutional Layer (<b>CNN-3</b>), CNN Model with Four Convolutional Layer (<b>CNN-4</b>), CNN Model with Five Convolutional Layer (<b>CNN-5</b>) and CNN Model with Six Convolutional Layer (<b>CNN-6</b>)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Liver) were taken from KMC Liver Dataset (Kasturba Medical College 2021) and Multi-organ Nucleus Dataset (Kumar et al. 2017, 2020)</p> <p><b>Parameters:</b> F1-Score and Jaccard Index (<b>Jl</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thus capable of precisely segmenting nuclei and further achieving promising results on both the KMC liver and Kumar datasets. The F1-Score measured for the KMC Liver Dataset was 83.59, JI was 72.06, and F1-Score measured for the Kumar Dataset was 81.363, JI was 68.883</p> <p><b>Limitations:</b> Due to the skip connection in the model architecture, the proposed model does not perform well</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Dense CRF loss Adam optimised for 80 epochs in initial training, and each round of self-training was employed. The proposed framework comprises two learning stages to perform the following: Stage 1: when a partially labelled nuclei location was provided, the same semi-supervised strategy was used to learn a detection model; Stage 2: when the nuclei location was detected, the same segmentation model was trained. The Voronoi label and cluster label were generated from the points detected</p> <p><b>Comparison:</b> CNN Model with Three Convolutional Layer (<b>CNN-3</b>) and DIST</p> <p><b>Datasets:</b> Cell images of one distinct organ (Lung) were taken from Lung Cancer (<b>LC</b>) Dataset (Kumar et al. 2017) and Multi-Organ (<b>MO</b>) Dataset (Kumar et al. 2017)</p> <p><b>Parameters:</b> Accuracy (<math>A_{GC}</math>), Object-level Dice coefficient (<math>Dice_{obj}</math>), Aggregated Jaccard Index (<b>AJI</b>) and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models; thereby further achieving competitive performance while requiring significantly less annotation effort. The measured values of <math>A_{CC}=0.9615</math>, <math>F1-Score=0.8771</math>, <math>Dice_{obj}=0.8521</math>, <math>AJI=0.6979</math>, and <math>ACC=0.9194</math>, <math>F1-Score=0.8100</math>, <math>Diceobj=0.6763</math>, and <math>AJI=0.3919</math> for the LC Dataset and MO Dataset, respectively</p> <p><b>Limitations:</b> The model utilizing cluster labels is unable to distinguish between nearby nuclei even when it has no access to Voronoi edge information</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned Adam's optimizer was employed. The DeepBC model identified and extracted the low-level and representative features that were considered best for the classifiers with discriminatory analysis by learning the hierarchical features in the training model</p> <p><b>Comparison:</b> GoogleNet, AlexNet, Residual Network (<b>ResNet</b>) and Visual Geometry Group with 16 Layer (<b>VGG-16</b>)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Breast Cancer Histopathological Dataset (<b>BreakHis dataset</b>) (Spanhol et al. 2015)</p> <p><b>Parameters:</b> Patient accuracy, Image accuracy and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby indicating that the model had favorable robustness and generalization, which could turn out to be advantageous for clinical classifications of breast cancer. Accuracy (Image)=96.43, Accuracy (Patient)=92.00, and F1-Score = 9.38 were measured</p> <p><b>Limitations:</b> The GoogleNet model's misclassification rate for fibroadenoma as well as lobular carcinoma is somehow lower than the DeepBC model</p>
<b>Year: 2020</b> Qu et al. (2020) proposed a weakly supervised segmentation framework based on partial points annotation in Histopathology Images	<p><b>Year: 2020</b> Wenzhong et al. (2020) proposed <b>DeepBC</b> for classifying the pathological images of breast cancer</p>

Table 1 (continued)

Paper details	
<b>Year: 2020</b> Ahamed et al. (2020) proposed an architecture, combining U-Net and Neural Ordinary Differential Equations for semantic segmentation in medical images	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Binary Cross-entropy (BCE) and Dice Loss Adam's optimizer was employed. The proposed model used the Ordinary Differential Equations (ODE) Block in a customised U-Net architecture</p> <p><b>Comparison:</b> Fully Convolutional Network (FCN), U-Net and Buda's U-Net</p> <p><b>Datasets:</b> Cell images of one distinct organ (Brain) were taken from Nuclei Images Dataset (Buda et al. 2019), Brain MRI Images Dataset (Buda 2020) and self-supervised dataset (Deng et al. 2009)</p> <p><b>Parameters:</b> Intersection Over Union (IOU) and Dice Coefficient (DC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby indicating that it used less memory and performed better with the same environmental setup for all the datasets. The measured values of IOU = 0.796, IoU = 0.948, and IOU = 0.276 were DC = 0.312 for the nuclei images dataset, the self-supervised dataset, and the brain MRI images dataset, respectively</p> <p><b>Limitations:</b> Due to the skip connection in the model architecture, the proposed model does not perform well</p>
<b>Year: 2020</b> Mehta et al. (2020) proposed a novel attention-based network: Holistic Attention Network (HATNet) for breast biopsy images classification	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Cross-entropy Loss</p> <p>An Adam optimizer with a learning rate warm-up strategy was employed. The number of parameters is 3.10 million. HATNet outperformed the popular approach known as "bag-of-words and further, for encoding global information, it made use of self-attention mechanisms without the help of any explicit supervision</p> <p><b>Comparison:</b> Pathologists, LAB &amp; LBP hand-crafted features (w/o saliency), LAB &amp; LBP hand-crafted features (w/ saliency), Bag-of-word (majority voting w/o saliency), Bag-of-word (majority voting w/ saliency), Bag-of-word (learned fusion w/o saliency), Bag-of-word (learned fusion w/saliency), MRSegNet with histogram and co-occurrence features, MRSegNet with structural features, Y-Net, HATNet (w/ ESPNet-v2), HATNet (w/ MobileNet-v2) and HATNet (w/ MNASNet)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Breast Biopsy Dataset (Elmore et al. 2015)</p> <p><b>Parameters:</b> Accuracy, F1-score, Sensitivity, Specificity and Receiver Operator Characteristic-Area Under Curve (ROC-AUC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby effectively aggregating inter-word and inter-bag representations. The measured values of accuracy are 0.71, F1-Score is 0.70, sensitivity is 0.70, specificity is 0.90, and ROC-AUC is 0.90</p> <p><b>Limitations:</b> The proposed model cannot perform well for blurry and complex images</p>
<b>Year: 2020</b> Celik et al. (2020) proposed an Invasive Ductal Carcinoma (IDC) detection task with histopathological images	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Not mentioned</p> <p>The deep learning pre-trained models Residual Network-50 (ResNet-50) and Densely Connected Convolutional Networks-161 (DenseNet-161) were used for the IDC detection task. The transfer learning method was used to design the model and figure out how to improve it</p> <p><b>Comparison:</b> Cruz-Roa et al. (2014), Janowczyk and Madabhushi (2016), Reza and Ma (2018) and Romero et al. (2019)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from IDC dataset (Janowczyk and Madabhushi 2016)</p> <p><b>Parameters:</b> F1-Score and Balanced Accuracy (BAC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance (classification accuracy) as compared to other models, thereby making it eligible to be tested on a much larger and more diverse dataset. The measured F1-Score is 94.11% and the BAC is 91.57%</p> <p><b>Limitations:</b> The images used in the test set are not utilised in the training set, and training is only applied to the last layers of the models</p>



Table 1 (continued)

Paper details	
<b>Year: 2020</b> Feng et al. (2020) proposed multiscale image processing method in histopathological images	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned</p> <p>The proposed method basically focused on pyramidal sampling, wherein U-Net was employed over seven layers of diverse resolutions. To overcome the difference that was generated in staining technique, a normalisation mechanism, namely a partial colour mechanism, was adopted. Further, to remove the unwanted seams between blocks, a weighted overlapping method was implemented</p> <p><b>Comparison:</b> Deeplab-v3, Graph Convolutional Network (GCN), Segmentation Network (SegNet), Deconvolutional Network (DeconvNet), Pyramid Scene Parsing with Segmentation Network (PSPNet), Attention U-Net and Fully Convolutional Network (FCN)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Liver) were taken from a dataset that comes from the 2019 MICCAI PAIP Challenge (PAIP2019 2019)</p> <p><b>Parameters:</b> F1 score, Jaccard similarity score and directed Hausdorff distance</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby achieving better scores as discrete partial color normalization and weighted overlapping techniques were used during pre-processing and prediction. The measured values of the F1-score are 0.465, the Jaccard score is 0.904, and the Hausdorff distance is 4.793</p> <p><b>Limitations:</b> The method should admit to some inaccuracy because not every layer of prediction is flawlessly accurate</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Perturbation-sensitive Sample Mining loss, Mask-guided Distillation loss and the combination of Classification, Regression and Segmentation loss Stochastic Gradient Descent (SGD) was employed as an optimizer. The proposed method was initially applied to approximate the sensitivity of method was initially applied to approximate the sensitivity of perturbations, thereby enabling the collection of useful samples from massive cases. An additional, predicted segmentation mask was adopted to eradicate the inescapable noise, particularly from the background region</p> <p><b>Comparison:</b> Instance Relation Network (IR-Net), Object Detection by Knowledge Distillation (ODKD) and Fine-grained Feature Limitation (FFI)</p> <p><b>Datasets:</b> The dataset was created containing liquid-based Pap test specimen of 82 different patients and imaged in <math>\times 40</math> resolutions with <math>0.2529 \mu\text{m}</math> per pixel</p> <p><b>Parameters:</b> Average Jaccard Index (AJI) and mean Average Precision (mAP)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby facilitating its further adaptation to be implemented over other semi-supervised medical image-based segmentation tasks. The measured values of AJI are 67.12% and 40.52%, respectively</p> <p><b>Limitations:</b> The approach's mAP score is not higher than the comparative ODKD method. It penalizes categorization and feature mismatch as the root cause in all areas. Thus, the introduction of noise is unavoidable</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned Adam's optimizer was employed. The proposed model was applied to diagnose disease by picking the discriminative instance, and the network designed for the same has the ability to seize instance-wise dependencies and recalibrate its features as per the importance of the coefficient attained from the fused features</p> <p><b>Comparison:</b> CNN-Vote-LR, CNN-Vote-SVM, MIMLNN, MI-NET-RC, CNN-Design Feat-RF, MI-NET-DS, MAXMIN-Layer, Attention-MIP and MISVM</p> <p><b>Datasets:</b> A Whole Slide Gastric Image (WSGI) named dataset was designed and created comprising of 608 whole slide images collected from different patients with image-level label comprising of three classes namely Normal, Dysplasia, and Cancer [Number represents the abnormal grading]</p> <p><b>Parameters:</b> Average Score and Accuracy</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby highlighting the essence of the proposed model as a recalibration module with the ability to automatically figure out the crucial instances for image-level prediction. However, due to the limited GPU model, the restriction lies in training the two-stage framework at once rather than separately. For the WSGI dataset, the average score is 0.923 and the accuracy is 86.5 percent</p> <p><b>Limitations:</b> Due to technological difficulties, the work's restriction is that the two-stage framework was trained independently (e.g., with limited GPU memory). In this sense, the localization network's characteristics for extracting instances might not be the ideal option for the RMDL network</p>
<b>Year: 2020</b> Zhou et al. (2020) proposed Mask-guided Mean Teacher framework with Perturbation-sensitive Sample Mining (MIMT-PSM)	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned</p> <p>The proposed model was better in terms of performance as compared to other models, thereby facilitating its further adaptation to be implemented over other semi-supervised medical image-based segmentation tasks. The measured values of AJI are 67.12% and 40.52%, respectively</p> <p><b>Limitations:</b> The approach's mAP score is not higher than the comparative ODKD method. It penalizes categorization and feature mismatch as the root cause in all areas. Thus, the introduction of noise is unavoidable</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned Adam's optimizer was employed. The proposed model was applied to diagnose disease by picking the discriminative instance, and the network designed for the same has the ability to seize instance-wise dependencies and recalibrate its features as per the importance of the coefficient attained from the fused features</p> <p><b>Comparison:</b> CNN-Vote-LR, CNN-Vote-SVM, MIMLNN, MI-NET-RC, CNN-Design Feat-RF, MI-NET-DS, MAXMIN-Layer, Attention-MIP and MISVM</p> <p><b>Datasets:</b> A Whole Slide Gastric Image (WSGI) named dataset was designed and created comprising of 608 whole slide images collected from different patients with image-level label comprising of three classes namely Normal, Dysplasia, and Cancer [Number represents the abnormal grading]</p> <p><b>Parameters:</b> Average Score and Accuracy</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby highlighting the essence of the proposed model as a recalibration module with the ability to automatically figure out the crucial instances for image-level prediction. However, due to the limited GPU model, the restriction lies in training the two-stage framework at once rather than separately. For the WSGI dataset, the average score is 0.923 and the accuracy is 86.5 percent</p> <p><b>Limitations:</b> Due to technological difficulties, the work's restriction is that the two-stage framework was trained independently (e.g., with limited GPU memory). In this sense, the localization network's characteristics for extracting instances might not be the ideal option for the RMDL network</p>
<b>Year: 2019</b> Wang et al. (2019) proposed Recalibrated Multi-instance Deep Learning method (RMDL) for Whole Slide Gastric Image Classification	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned</p> <p>The proposed model was better in terms of performance as compared to other models, thereby facilitating its further adaptation to be implemented over other semi-supervised medical image-based segmentation tasks. The measured values of AJI are 67.12% and 40.52%, respectively</p> <p><b>Limitations:</b> The approach's mAP score is not higher than the comparative ODKD method. It penalizes categorization and feature mismatch as the root cause in all areas. Thus, the introduction of noise is unavoidable</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned Adam's optimizer was employed. The proposed model was applied to diagnose disease by picking the discriminative instance, and the network designed for the same has the ability to seize instance-wise dependencies and recalibrate its features as per the importance of the coefficient attained from the fused features</p> <p><b>Comparison:</b> CNN-Vote-LR, CNN-Vote-SVM, MIMLNN, MI-NET-RC, CNN-Design Feat-RF, MI-NET-DS, MAXMIN-Layer, Attention-MIP and MISVM</p> <p><b>Datasets:</b> A Whole Slide Gastric Image (WSGI) named dataset was designed and created comprising of 608 whole slide images collected from different patients with image-level label comprising of three classes namely Normal, Dysplasia, and Cancer [Number represents the abnormal grading]</p> <p><b>Parameters:</b> Average Score and Accuracy</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby highlighting the essence of the proposed model as a recalibration module with the ability to automatically figure out the crucial instances for image-level prediction. However, due to the limited GPU model, the restriction lies in training the two-stage framework at once rather than separately. For the WSGI dataset, the average score is 0.923 and the accuracy is 86.5 percent</p> <p><b>Limitations:</b> Due to technological difficulties, the work's restriction is that the two-stage framework was trained independently (e.g., with limited GPU memory). In this sense, the localization network's characteristics for extracting instances might not be the ideal option for the RMDL network</p>

Table 1 (continued)

Paper details	
<b>Year: 2020</b> Jha et al. (2020) proposed <b>Double U-Net</b> for Medical Image Segmentation	<p><b>Backbone:</b> Visual Geometry Group Network (<b>VGG-19</b>)</p> <p><b>Loss:</b> Binary Cross-entropy (BCE)</p> <p>Adam's optimizer was employed. The proposed model introduced yet another U-Net at the bottom of the network for efficient capture of supplementary semantic information. Further, Atrous Spatial Pyramid Pooling (ASPP) was adapted to capture contextual data within the network. Additionally, post-processing techniques such as conditional random field and Otsu threshold can improve the result significantly</p> <p><b>Comparison:</b></p> <ul style="list-style-type: none"> <li>(i) <b>Evaluated on the 2015 MICCAI sub-challenge on automatic polyp detection dataset</b> – Fully Convolutional Network—Visual Geometry Group Network (<b>FCN-VGG</b>), Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>) with ResNet-101 and U-Net</li> <li>(ii) <b>Evaluated on the CVC-ClinicDB dataset</b> – Fully Convolutional Network (<b>FCN</b>), CNN, Segmentation Network (<b>SegNet</b>), Multi-scale patch-based CNN, MultiResUNet with data augmentation, Conditional generative adversarial network and U-Net</li> <li>(iii) <b>Evaluated on the Lesion Boundary Segmentation challenge dataset</b> – U-Net and Multi-ResUNet</li> <li>(iv) <b>Evaluated on the 2018 Data Science Bowl Challenge dataset</b> – U-Net and UNet++</li> </ul> <p><b>Datasets:</b> Cell images of one distinct organ (Colon) was taken from 2015 MICCAI sub-challenge on automatic polyp detection Dataset (Bernal et al. 2017), CVC-ClinicDB Dataset (Mehta et al. 2020), Lesion Boundary Segmentation Challenge Dataset (Codella et al. 2019; Tschandl et al. 2018), 2018 Data Science Bowl Challenge Dataset (Caicedo et al. 2019; Data science bowl 2018)</p> <p><b>Parameters:</b> Intersection over Union (<b>IoU</b>), Precision and Recall</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby highlighting the fact that usage of Double U-Net could act as a sturdy baseline in cases of both medical image segmentation and cross-dataset evaluation testing, enhancing the generalizability of Deep Learning (DL) models. The measured values of (DSC = 0.7649, mIoU = 0.6255, Recall = 0.7156 and Precision = 0.8007), (DSC = 0.9239, mIoU = 0.8611, Recall = 0.8457 and Precision = 0.9592), (DSC = 0.8962, mIoU = 0.8212, Recall = 0.8780 and Precision = 0.9459) and (DSC = 0.9133, mIoU = 0.8407, Recall = 0.6407 and Precision = 0.9496) for 2015 MICCAI sub-challenge on automatic polyp detection, CVC-ClinicDB, Lesion Boundary Segmentation challenge and 2018 Data Science Bowl Challenge Dataset, respectively</p> <p><b>Limitations:</b> One drawback of the Double U-Net is that it employs more parameters than the U-Net, which extends training time</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Binary Cross-entropy (<b>BCE</b>) and Dice Coefficient Loss Adam's optimizer was employed. Two stages were adopted for segmentation: H&amp;E-stained images were pre-processed to reduce the variance; a further second stage takes in the output of stage 1 as input given to the LinkNET network, which comprises down-sampling as well as up-sampling layers</p> <p><b>Comparison:</b> Deep Convolutional Neural Network (<b>Deep CNN</b>), U-Net</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Benchmark Dataset (Data science bowl 2018; Menze et al. 2014)</p> <p><b>Parameters:</b> Intersection over Union (<b>IoU</b>), Dice Coefficient (<b>DC</b>), Accuracy</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance because all extra features were mined through up-sampling and it was capable of easily localizing and learning the representation. The measured values of the benchmark dataset are IoU = 89.8, DC = 0.89, and accuracy = 97.2</p> <p><b>Limitations:</b> Due to the skip connection in the model architecture, the proposed model does not perform well</p>
<b>Year: 2020</b> Natarajan et al. (2020) utilised <b>LinkNET-34</b> architecture for semantic segmentation of nuclei from the H&E-stained breast cancer histopathology images	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Binary Cross-entropy (<b>BCE</b>) and Dice Coefficient Loss</p> <p>Adam's optimizer was employed. Two stages were adopted for segmentation: H&amp;E-stained images were pre-processed to reduce the variance; a further second stage takes in the output of stage 1 as input given to the LinkNET network, which comprises down-sampling as well as up-sampling layers</p> <p><b>Comparison:</b> Deep Convolutional Neural Network (<b>Deep CNN</b>), U-Net</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Benchmark Dataset (Data science bowl 2018; Menze et al. 2014)</p> <p><b>Parameters:</b> Intersection over Union (<b>IoU</b>), Dice Coefficient (<b>DC</b>), Accuracy</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance because all extra features were mined through up-sampling and it was capable of easily localizing and learning the representation. The measured values of the benchmark dataset are IoU = 89.8, DC = 0.89, and accuracy = 97.2</p> <p><b>Limitations:</b> Due to the skip connection in the model architecture, the proposed model does not perform well</p>

Table 1 (continued)

Paper details	
<b>Year:</b> 2020	<b>Backbone:</b> Not mentioned
Wang et al. (2020) proposed <b>Bending Loss Regularized Network</b> to tackle the challenge of segmenting overlapped nuclei in histopathology images	<b>Loss:</b> Bending loss The proposed method applied high and low penalties to contour points with large and small curvature, respectively. In general, the bending loss was reduced to produce nuclei boundary points with smooth curves, avoiding the generation of boundaries for two or more nuclei that were touching
<b>Comparison:</b>	Fully Convolutional Neural Network ( <b>FCN-8</b> ), U-Net, Segmentation Network ( <b>SegNet</b> ), Deep Convolutional Auto-encoder Network ( <b>DCAN</b> ), DIST and Horizontal and Vertical Distance Network ( <b>HoVer-Net</b> )
<b>Datasets:</b>	Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from Multi-Organ Nuclei Segmentation (MoNuSeg) Dataset (Kumar et al. 2017, 2020)
<b>Parameters:</b>	Dice, Aggregated Jaccard Index ( <b>AJI</b> ), Recognition Quality ( <b>RQ</b> ), Segmentation Quality ( <b>SQ</b> ), Panoptic Quality ( <b>PQ</b> )
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models. Thus enhancing the model and thereby enabling it to be applied further to other deep learning-based segmentation tasks. The measured values of the MoNuSeg Dataset are: $AJI=0.621$ , $Dice=0.813$ , $RQ=0.781$ , $SQ=0.762$ , and $PQ=0.596$ for the same organ test, and $AJI=0.641$ , $Dice=0.837$ , $RQ=0.760$ , $SQ=0.775$ and $PQ=0.592$ for the different organ test
<b>Limitations:</b>	Some histopathological images still pose a problem for the segmentation of overlapping nuclei, and the SQ results for different organ tests aren't as good as the DIST model
<b>Year:</b> 2021	<b>Backbone:</b> Visual Geometric Group ( <b>VGG</b> )
Chanchal et al. (2021b) proposed Separable Convolutional Pyramid Pooling Network ( <b>SCPP-Net</b> ) for segmentation of kidney and breast histopathology images	<b>Loss:</b> Binary Cross-entropy ( <b>BCE</b> ) Adam optimizer was employed, with the number of parameters being 5,088,955. The proposed unit emphasized two significant facts: keeping the kernel size fixed and increasing the corresponding fields by varying the four dilation rates; and reducing the trainable parameters by means of depth-wise separable convolution
<b>Comparison:</b>	U-Net, SegNet, Attention U-Net, DIST and Atrous Spatial Pyramid Pooling U-Net ( <b>ASPP U-Net</b> )
<b>Datasets:</b>	Cell images of two distinct organs (Kidney and Breast) were taken from Haematoxylin and Eosin (H&E)-stained Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018), H&E-stained Kidney Dataset (Irshad et al. 2015) and Multiple Organs Multi-Disease Histopathology Dataset (Kumar et al. 2017, 2020)
<b>Parameters:</b>	F1-Score and Aggregated Jaccard Index ( <b>AJI</b> )
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, and in doing so, two significant flaws were overcome by the proposed model, i.e., separating nuclei from complex, structured histopathology images with varying histology and molecular characteristics and reducing the computational complexity and total trainable parameters. The measured values of (F1-Score = 0.9203, $AJI=0.8592$ ), (F1-Score = 0.8168, $AJI=0.6998$ ), and (F1-Score = 0.8010, $AJI=0.6710$ ) for the Kidney, TNBC, and Multiple Organs Multi-Disease Histopathology Datasets, respectively
<b>Limitations:</b>	The distinction of eosin cells from hematoxylin-stained nuclei was the only focus of the study's efforts. Pathologists believe that cytoplasmic eosin should be tested in cases of higher-grade malignancy. The segmentation of overlapping nuclei is another restriction that hasn't been fully overcome

Table 1 (continued)

Paper details	
<b>Year:</b> 2021	<b>Backbone:</b> U-Net
Chen et al. (2021) proposed Context-aware Polygon Proposal Network (CPP-Net) for nucleus segmentation	<b>Loss:</b> Shape-Aware Perceptual (SAP) Adam's optimizer was employed. The proposed model, rather than sampling a single pixel, tended to sample a point set, thus significantly enhancing the contextual information and refining the sturdiness of prediction. Further, a confidence-based weighting module and shape-aware perceptual (SAP) loss were employed that adaptively fused the predictions from the sampled point set and constrained the shape of the predicted polygons, respectively
<b>Comparison:</b>	(i) <b>Evaluated on the DSB2018 Dataset –</b> Mask Regional Convolutional Neural Network (Mask R-CNN), Object Detection with Star-convex Shapes (StarDist), Keypoint Graph, Horizontal and Vertical Distances Network (HoVer-Net) and PatchPerPix (ii) <b>Evaluated on the BBBC006 Dataset –</b> Instance Embedding, Keypoint Graph, Horizontal and Vertical Distances Network (HoVer-Net) and Object Detection with Star-convex Shapes (StarDist) (iii) <b>Evaluated on the PanNuke Dataset -</b> Mask Regional Convolutional Neural Network (Mask R-CNN), Micro-Net, Horizontal and Vertical Distances Network (HoVer-Net) and Object Detection with Star-convex Shapes (StarDist)
<b>Datasets:</b>	Cell images of two distinct organs (Breast and lung) were taken from Data Science Bowl 2018 (DSB2018) Dataset (Caicedo et al. 2019; Data science bowl 2018), Broad Bioimage Benchmark Collection (BBBC006) Dataset (Ljosa et al. 2012) and PanNuke Dataset (Gamper et al. 2019, 2020)
<b>Parameters:</b>	Mean Average Precision (Mean AP), multi-class Panoptic Quality (mPQ), and binary Panoptic Quality (bPQ)
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thereby requiring extensive ablation studies to justify the effectiveness of each proposed component. The measured values of Mean AP are 0.7086 and 0.6804 for the DSB2018 and BBBC006 datasets, respectively. And for the PanNuke Dataset, the measured value of mPQ is 0.4817 and bPQ is 0.6767
<b>Limitations:</b>	Lower IoU values for the APs mean that SAP loss does not significantly enhance performance. Due to this phenomenon, the prediction error in nucleus form is penalized by the SAP loss more than the localization or detection mistakes
<b>Features:</b>	<b>Backbone:</b> U-Net <b>Loss:</b> Weighted Cross-entropy (WCE) and Dice Loss Adam's optimizer was employed. The proposed method basically aim to identify whether deep convolutional neural networks would be suitable enough to train with transfer learning, on a set of histopathological images independent of breast tissue to segment tumour nuclei of the breast or not
<b>Comparison:</b>	(i) <b>Evaluated on the MoNuSeg Dataset -</b> Otsu, Watershed, Fiji, U-Net like Deep Convolutional Neural Networks (U-Net like DCNNs), Mask Regional Convolutional Neural Network (Mask R-CNN) and U-Net Ensemble (ii) <b>Evaluated on the TNBC Dataset -</b> U-Net like Deep Convolutional Neural Networks (U-Net like DCNNs), Mask Regional Convolutional Neural Network (Mask R-CNN) and U-Net Ensemble
<b>Datasets:</b>	Cell images of one distinct organ (Breast) was taken from Multi-Organ Nucleus Segmentation (MoNuSeg) Dataset (Kumar et al. 2017, 2020), Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)
<b>Parameters:</b>	Aggregated Jaccard Index (AJI) and mean Average Precision (mAP)
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thereby indicating that tumour nuclei in the breast could be accurately segmented. the measured values of AJI = 0.53, mAP = 0.39, and AJI = 0.54, mAP = 0.54, for the MoNuSeg and TNBC datasets, respectively
<b>Limitations:</b>	The very small dataset used for training and testing, based on the few open-source datasets, is a weakness of the work
<b>Year:</b> 2021	<b>Backbone:</b> U-Net
Lagree et al. (2021) proposed a Gradient Boosting U-Net (GB U-Net) for breast tumour cell nuclei segmentation	<b>Loss:</b> Weighted Cross-entropy (WCE) and Dice Loss Adam's optimizer was employed. The proposed method basically aim to identify whether deep convolutional neural networks would be suitable enough to train with transfer learning, on a set of histopathological images independent of breast tissue to segment tumour nuclei of the breast or not
<b>Comparison:</b>	(i) <b>Evaluated on the MoNuSeg Dataset -</b> Otsu, Watershed, Fiji, U-Net like Deep Convolutional Neural Networks (U-Net like DCNNs), Mask Regional Convolutional Neural Network (Mask R-CNN) and U-Net Ensemble (ii) <b>Evaluated on the TNBC Dataset -</b> U-Net like Deep Convolutional Neural Networks (U-Net like DCNNs), Mask Regional Convolutional Neural Network (Mask R-CNN) and U-Net Ensemble
<b>Datasets:</b>	Cell images of one distinct organ (Breast) was taken from Multi-Organ Nucleus Segmentation (MoNuSeg) Dataset (Kumar et al. 2017, 2020), Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)
<b>Parameters:</b>	Aggregated Jaccard Index (AJI) and mean Average Precision (mAP)
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thereby indicating that tumour nuclei in the breast could be accurately segmented. the measured values of AJI = 0.53, mAP = 0.39, and AJI = 0.54, mAP = 0.54, for the MoNuSeg and TNBC datasets, respectively
<b>Limitations:</b>	The very small dataset used for training and testing, based on the few open-source datasets, is a weakness of the work

Table 1 (continued)

Paper details	
<b>Year: 2021</b> Camalan et al. (2021) proposed a Deep Convolutional Neural Network (D-CNN) oral lesion classification system for clinical oral photographic images	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Cross-entropy Stochastic Gradient Descent (SGD) was used as an optimizer. The proposed method aimed to categorize images under headings such as "suspicious" and "normal" by finalizing transfer learning over Inception-ResNet-V2 and further generating automated heat maps to focus on the region of the images presumably involved in decision-making. Transfer learning was performed on a limited number of image samples for oral dysplasia</p> <p><b>Comparison:</b> Inception ResNet-V2, Inception V3, Visual Geometry Group with 16 Layers (VGG-16) and ResNet-101</p> <p><b>Datasets:</b> Cell images of one distinct organ (oral) were taken from Sheffield (UK) Dataset (Ethical approval for the Sheffield cohort was obtained for this study from the HRA and Health and Care Research Wales (HCRW) 2018), Piracicaba (Brazil) Dataset (Piracicaba Dental Ethical Committee 2019)</p> <p><b>Parameters:</b> Accuracy, F1-Score, Precision, and Recall</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models; however, the number of data points was sufficient yet small (54), making it difficult to train and test the system. Further, to overcome the limitation, the number of images was increased by augmenting and splitting the images into patches. The measured values of (Accuracy (%) = 73.6 ± 19, F1-Score (%) = 97.9, Precision (%) = 95.4, Recall (%) = 100.0) and (Accuracy (%) = 90.9 ± 12, F1-Score (%) = 87.2, Precision (%) = 99.3, Recall (%) = 81.1) for Sheffield and Piracicaba Datasets, respectively</p> <p><b>Limitations:</b> There are some limitations to the study. The system was trained and tested on a total of 54 patients, which is a tiny sample size for independent testing, but it was enough to show that the methods worked. Another drawback is the process of choosing the patches. A percentage threshold for the total number of pixels was used to determine whether a patch is questionable or not</p>
<b>Year: 2021</b> Xiao et al. (2021) proposed a Polar representation-based nucleus segmentation model in non-small lung cancer histopathological images	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Polar centeredness and polar IoU loss Stochastic Gradient Descent (SGD) was used as an optimizer. The proposed module employed centre classification and length regression to produce the contour of the nucleus in a polar coordinate, thus playing a dynamic role</p> <p><b>Comparison:</b> U-Net, ExtremeNet, TensorMask, and PolarMask</p> <p><b>Datasets:</b> They were manually collected 4792 histopathological slides with the lesions caused by non-small cell lung cancer from Shandong Provincial Hospital</p> <p><b>Parameters:</b> F1-Score, Dice, Hausdorff and Aggregated Jaccard Index (AJI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thus indicating that the proposed approach could be a hypothetically valuable tool for clinical practices. The measured values of F1-score are: 0.8382 Dice=0.8092, Hausdorff = 11.2873, and AJI = 0.6873</p> <p><b>Limitations:</b> When the images have a complex shape and structure, segmentation becomes a difficult task for the model</p>
<b>Year: 2021</b> Jahamifar et. al. (2021) proposed an interactive semantic segmentation model for robust tissue region annotation in a semi-automated manner	<p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Soft Dice and Binary Cross-entropy (BCE) The proposed module used the Efficient-UNet architecture to improve recognition of regions of varying sizes and to train the model's four techniques to extract minimalistic and human-drawn-like guiding signals from GT masks. Smoothing filters were applied over the original mask to lessen the boundary noise and achieve an improved morphological skeleton</p> <p><b>Comparison:</b> U-Net, DeepLab v3, Efficient UNet-B0, Efficient UNet-B1 and Efficient UNet-B2</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Amgad dataset (Amgad et al. 2019)</p> <p><b>Parameters:</b> Dice, Accuracy (ACC) and Area Under Curve (AUC)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thus highlighting that the proposed module accelerated the interactive annotation process and outdid the existing automatic and interactive region segmentation models. The measured values of dice, accuracy, and AUC were 0.875, 0.984, and 0.995, respectively</p> <p><b>Limitations:</b> Lack of training data is one of the key causes of the lack of significant performance improvement as network scale increases. The network's architecture not being scalable is another issue</p>



Table 1 (continued)

Paper details	
<b>Year:</b> 2021	<b>Backbone:</b> Not Mentioned
Ali et al. (2021) proposed Pyramid Pooling U-Net (PPU-Net) for Nucleus Segmentation from Brightfield Cell Microscopy Images	<b>Loss:</b> Binary Cross-entropy (BCE) The Adam optimizer was used with a 1.3 million trainable parameter architecture. The proposed model post-processed the image probability maps they produced for evaluating the models and additionally tuned the model using images from the target domain (another data set) to train the source domain (one data set)
	<b>Comparison:</b> U-Net, U-Net++, DeepLab v3+, and Tiramisu
	<b>Datasets:</b> Cell images of three distinct organs (Kidney, Breast and Prostate) were taken from seven cells Line Dataset (Fishman et al. 2019) and LNCaP Dataset (Fishman et al. 2019)
	<b>Parameters:</b> F1-Score and Intersection over Union (IoU)
	<b>Inference:</b> The proposed model was better in terms of performance as compared to other models, as it employed simple light residual connections and attained decent outcomes in both bright field and fluorescence modalities. Further, the proposed model seemed extremely appropriate to be employed in large-scale experiments. For the Seven Cell Line Dataset, the measured values of F1 are 0.865; for the LNCaP Dataset, the measured values of F1 are 0.874. PPU-Net similarly
	<b>Limitations:</b> To determine the optimal approach in cell microscopy, specialized object delineation methods and unique data sets outside the purview of the work are required
<b>Year:</b> 2021	<b>Backbone:</b> Not mentioned
Ioannidis et al. (2021) proposed a Pathomics and Deep Learning methodology for Fluorescence Histology Images	<b>Loss:</b> Not mentioned The proposed module calculated the area of each object using the label function from Mahot's library, and it used all of the Pyradiomics library's accessible features, including statistical and higher-order statistical texture features. In addition, on the training set, three labeled nuclei types were achieved using the PyMRR library and the Mutual Information Differences (MID) technique to recognize an eloquent feature group
	<b>Comparison:</b> Xception, Residual Network (ResNet), Visual Geometry Group (VGG), Mobile-Net, Dense-Net and Nas-Net
	<b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Fluorescence Dataset (Ktomp et al. 2020)
	<b>Parameters:</b> Area Under Curve (AUC) and Accuracy (ACC)
	<b>Inference:</b> The proposed model was better in terms of performance as compared to other models in spite of the heterogeneity owing to the nonexistence of standardized image acquisition protocol, thus proving it to be applicable as a generalized and robust model in different scenarios. The measured values of AUC = 0.986 ± 0.033, ACC = 0.957 ± 0.105
	<b>Limitations:</b> The dataset employed (N = 79 photos) was quite small, which is where the constraint comes from. But the size of the dataset was carefully considered when deciding on the analytical pipeline. This was the main reason why more traditional methods were used

Table 1 (continued)

Paper details	
<b>Year: 2021</b>	<b>Backbone:</b> Not mentioned
Mahmood et al. (2021) proposed AI-based nuclear segmentation technique and Residual Skip Connections-based segmentation Network for Nuclei (R-SNN) of Nuclear Regions with Multi-Organ Histopathology Image	<b>Loss:</b> Cross-entropy Adam optimizer was employed, and R-SNN comprised 15,279,174 trainable parameters. The proposed method based on R-SNN used the stain normalization technique to come up with images with standard appearances and reduce the image's color variants and intensity. ResNet was proposed in the work basically to lessen information loss and resolve the vanishing gradient problem
<b>Year: 2021</b>	<b>Comparison:</b>
Hassan et al. (2021b) proposed cell nuclei segmentation method based on deep learning	(i) Evaluated on the TCGA Dataset - Cell Profiler (CP), Kumar et al. (2017), Kang et al. (2019a), Mahbod et al. (2019), Zeng et al. (2019) and Chidester et al. (2019) (ii) Evaluated on the TNBC Dataset - PangNet, Deconvolutional Network (DeconvNet), Fully Convolutional Network (FCN), Ensemble and Kang et al. (2019)
	<b>Datasets:</b>
	Cell images of nine distinct organs (Brain, Breast, Kidney, Liver, Prostate, Bladder, Colon, Stomach and Lungs) were taken from The Cancer Genome Atlas (TCGA) (Tomczak et al. 2015); The Cancer Genome Atlas (TCGA) 2016), Triple-Negative Breast Cancer (TNBC) (Naylor et al. 2018)
	<b>Parameters:</b>
	F1-measure, Dice's coefficient (DC), Aggregated Jaccard Index (AJI), Precision and Recall
	<b>Inference:</b>
	The proposed model was better in terms of performance (segmentation) as compared to other models without requiring additional sub-patch conversion. On the other hand, the proposed module required no post-processing, making it more appropriate and desirable. For the TCGA Dataset, the measured values of AJI are 0.6794, DC is 0.8084, and F1-measure is 0.8547. On the TNBC dataset, the measured values were AJI = 0.7332, DC = 0.8441, Precision = 0.8352, Recall = 0.8306, and F1-measure = 0.8329
	<b>Limitations:</b>
	The work has several restrictions. Stain normalization, a labor-intensive procedure that could potentially result in significant computer complexity, was first carried out. Second, it was challenging to distinguish nearby nuclei; they were treated as a single item. Third, many apps use whole-slide photos, which have dimensions far greater than 1000×1000 pixels
<b>Year: 2021</b>	<b>Backbone:</b> Not mentioned
Sohail et al. (2021b) proposed cell nuclei segmentation method based on deep learning	<b>Loss:</b> Cross-entropy Stochastic Gradient Descent (SGD) was used as an optimizer. The proposed work is intended to develop GUI-based cell nuclei segmentation software consisting of the proposed method along with other deep learning-based cell nuclei segmentation methods. The stain normalization method was applied to the proposed model, converting all images to its color space, and only one image was selected as the target
	<b>Comparison:</b>
	Fully Convolutional Densely Connected Convolutional Network-103 (FC DenseNet-103), Pyramid Scene Parsing with Segmentation Network (PSPSeg-Net), Self-Correction and Learnable Aggregation Network (LANet) w/o concat RGB
	<b>Datasets:</b>
	Cell images of seven distinct organs (Colon, Breast, Stomach, Kidney, Prostate, Bladder and Liver) were taken from Multi-Organ Nuclei Segmentation (MoNuSeg) Dataset (Kumar et al. 2017, 2020)
	<b>Parameters:</b>
	F1-Score and Aggregated Jaccard Index (AJI)
	<b>Inference:</b>
	The proposed model was better in terms of performance as compared to other models. The measured values of F1-Score and AJI for the MoNuSeg Dataset are 73.02% and 89.25%, respectively
	<b>Limitations:</b>
	A total of 30 WSI images were used to base all of the results, which is a very small sample size for independently developing and testing the system
<b>Year: 2021</b>	<b>Backbone:</b> Not mentioned
Sohail et al. (2021) multi-phase mitosis detection framework "MP-MitDet" for mitotic nuclei identification in breast cancer histopathological images	<b>Loss:</b> Multi-objective (Classification Loss, Box Regression Loss & Segmentation Loss) The workflow of the proposed model consisted of: automatic label-refiner representing weak labels with semi-semantic information; tissue-level mitotic region selection; blob analysis; and cell-level refinement. Trained Mask R-CNN was further used for applying pixel-by-pixel masks for the training set of the mitosis detection module
	<b>Comparison:</b>
	Wahab et al. (2019), Li et al. (2019b), Paeng et al. (2017), Akram et al. (2018) and Mahmood et al. (2020)
	<b>Datasets:</b>
	Cell images of one distinct organ (Breast) were taken from TUPAC16 Dataset (Akram et al. 2018)
	<b>Parameters:</b>
	F1-Score, Precision and Recall
	<b>Inference:</b>
	The proposed model was better in terms of performance as compared to other models, thereby enhancing the results by significantly reducing the mitosis-like-hard example with a significant precision (0.70 for custom, 0.73 for ensemble-based CNN) on the test dataset. For the TUPAC16 Dataset, the measured values of F1-Score are 0.75, Precision is 0.71, and Recall is 0.76
	<b>Limitations:</b>
	According to the F-score performance study of test images, traditional classifiers do not have sufficient representational ability to distinguish between two classes. The complexity of the data was probably the cause of the low performance

Table 1 (continued)

Paper details	
<b>Year: 2021</b> Shuvo et al. (2021) proposed Classifier and Localizer U-Net (CNL U-Net) for 2D multimodal Biomedical Image Segmentation	<p><b>Backbone:</b> U-Shaped Network</p> <p><b>Loss:</b> Dice Loss, Binary Cross-entropy (BCE) and Mean Squared Error (MSE) Adam optimizer to train the models and minimise the loss function was employed, and the model had 11.5 million parameters. The proposed module consisted of a pre-trained encoder supplemented with transfer learning techniques for sufficient learning even from a small amount of existing data. Furthermore, the proposed model would use modified skip connections to reduce semantic gaps between encoder-decoder layer levels</p> <p><b>Comparison:</b> U-Net and Segmentation Network (SegNet)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Brain) was taken from Shenzhen and Montgomery Country (MC) Dataset or Chest X-Ray Image Dataset (Jaeger et al. 2013; Candemir et al. 2013), ISIC 2018-Skin Lesion Segmentation Competition Dataset or Dermoscopy Dataset (Codella et al. 2019; Tschandl et al. 2018), 2018 Data Science Bowl Competition Dataset or Microscopy dataset (Caicedo et al. 2019; Data science bowl 2018), Ultrasound Nerve Segmentation Dataset (Ultrasound nerve segmentation 2016) and LGG Segmentation Dataset or Brain MRI Image Dataset (Buda 2020; Buda et al. 2019)</p> <p><b>Parameters:</b> Dice Coefficient (DC) and jaccard Index (JI)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby demonstrating its ability, particularly in the case of detailed segmentation boundaries, and thus reducing false positives and negatives with faster learning ability. The measured values of (DC=0.9594, JI=0.9226), (DC=0.8421, JI=0.7359), (DC=0.9282, JI=0.8668), (DC=0.7647, JI=0.6209), and (DC=0.7417, JI=0.5911) for the chest X-ray image dataset, dermoscopy dataset, microscopy dataset, ultrasound dataset, and brain MRI image dataset, respectively</p> <p><b>Limitations:</b> The CNL module's classification performance in the model is weak</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned</p> <p>The proposed model consisted of attention mechanisms (AM) to extract attention areas and image classification (IC) modules. Once the attention areas were extracted using the AM module, these regions were used to train the CNN model. Using the probability-based ensemble learning (EL) classification technique, image level was obtained from patch level output of CNN</p> <p><b>Comparison:</b> Level-Set, Dense Conditional Random Fields (Dense CRF), U-Net, Watershed, Markov Random Field (MRF), Otsu and Seg-Net</p> <p><b>Datasets:</b> Cell images of one distinct organ (Stomach) were taken from Haematoxylin and Eosin (H&amp;E) stained Gastric Histopathological Image Dataset (Zhang et al. 2018)</p> <p><b>Parameters:</b> Accuracy, Sensitivity, Specificity, Precision and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby paving the way toward a human-machine collaboration pattern for early diagnosis of gastric cancer. For the dataset, the measured values are: accuracy=0.914, sensitivity=0.757, specificity=0.967, precision=0.883, and F1-score=0.815</p> <p><b>Limitations:</b> For the proposed model, segmentation is a challenging task when the images have a complex shape and structure</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Cross-entropy Stochastic Gradient Descent (SGD) was used as an optimizer</p> <p>Three frameworks were proposed in the proposed method: an image-level framework, a cell-level framework, and a combination framework. Herein, the CNN model was trained to forecast the cancerous probability of the tissue, to analyze the WSIs in high magnification, and to further combine feature vectors from the image-level framework with heat map features from the cell-level framework for additional classification</p> <p><b>Comparison:</b> Multiple Instance Learning with Deep Learning Approach (MIL-CNN) and Deep neural network</p> <p><b>Datasets:</b> Cell images of one distinct organ (Colon) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Accuracy, Precision, Recall and F1-Score</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance because it introduced a superior combination framework, allowing it to achieve WSI classification and localization while utilizing only global labels. For the TCGA Dataset, the measured values of accuracy were 0.946</p> <p><b>Limitations:</b> The TCGA WSIs lack sufficient light blood cells due to technological and procedural constraints. Also, even after being trained, the cell-level model couldn't figure out these things about blood cells</p>
<b>Year: 2021</b> Li et al. (2021a) proposed a Hierarchical Conditional Random Field based Attention Mechanism (HCRF-AM) over Gastric Histopathology Image Classification (GHIC) tasks	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Not mentioned</p> <p>The proposed model consisted of attention mechanisms (AM) to extract attention areas and image classification (IC) modules. Once the attention areas were extracted using the AM module, these regions were used to train the CNN model. Using the probability-based ensemble learning (EL) classification technique, image level was obtained from patch level output of CNN</p> <p><b>Comparison:</b> Level-Set, Dense Conditional Random Fields (Dense CRF), U-Net, Watershed, Markov Random Field (MRF), Otsu and Seg-Net</p> <p><b>Datasets:</b> Cell images of one distinct organ (Stomach) were taken from Haematoxylin and Eosin (H&amp;E) stained Gastric Histopathological Image Dataset (Zhang et al. 2018)</p> <p><b>Parameters:</b> Accuracy, Sensitivity, Specificity, Precision and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby paving the way toward a human-machine collaboration pattern for early diagnosis of gastric cancer. For the dataset, the measured values are: accuracy=0.914, sensitivity=0.757, specificity=0.967, precision=0.883, and F1-score=0.815</p> <p><b>Limitations:</b> For the proposed model, segmentation is a challenging task when the images have a complex shape and structure</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Cross-entropy Stochastic Gradient Descent (SGD) was used as an optimizer</p> <p>Three frameworks were proposed in the proposed method: an image-level framework, a cell-level framework, and a combination framework. Herein, the CNN model was trained to forecast the cancerous probability of the tissue, to analyze the WSIs in high magnification, and to further combine feature vectors from the image-level framework with heat map features from the cell-level framework for additional classification</p> <p><b>Comparison:</b> Multiple Instance Learning with Deep Learning Approach (MIL-CNN) and Deep neural network</p> <p><b>Datasets:</b> Cell images of one distinct organ (Colon) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Accuracy, Precision, Recall and F1-Score</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance because it introduced a superior combination framework, allowing it to achieve WSI classification and localization while utilizing only global labels. For the TCGA Dataset, the measured values of accuracy were 0.946</p> <p><b>Limitations:</b> The TCGA WSIs lack sufficient light blood cells due to technological and procedural constraints. Also, even after being trained, the cell-level model couldn't figure out these things about blood cells</p>
<b>Year: 2021</b> Zhou et al. (2021) proposed Three frameworks for Histopathology classification and localization of colorectal cancer	<p><b>Backbone:</b> Not mentioned</p> <p><b>Loss:</b> Cross-entropy Stochastic Gradient Descent (SGD) was used as an optimizer</p> <p>Three frameworks were proposed in the proposed method: an image-level framework, a cell-level framework, and a combination framework. Herein, the CNN model was trained to forecast the cancerous probability of the tissue, to analyze the WSIs in high magnification, and to further combine feature vectors from the image-level framework with heat map features from the cell-level framework for additional classification</p> <p><b>Comparison:</b> Multiple Instance Learning with Deep Learning Approach (MIL-CNN) and Deep neural network</p> <p><b>Datasets:</b> Cell images of one distinct organ (Colon) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016)</p> <p><b>Parameters:</b> Accuracy, Precision, Recall and F1-Score</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance because it introduced a superior combination framework, allowing it to achieve WSI classification and localization while utilizing only global labels. For the TCGA Dataset, the measured values of accuracy were 0.946</p> <p><b>Limitations:</b> The TCGA WSIs lack sufficient light blood cells due to technological and procedural constraints. Also, even after being trained, the cell-level model couldn't figure out these things about blood cells</p>

Table 1 (continued)

Paper details	
<b>Year: 2021</b> Podder et al. (2021) proposed Mask Regional Convolutional Neural Network ( <b>Mask R-CNN</b> ) for detection of COVID-19 using chest X-ray images	<p><b>Features:</b> <b>Backbone:</b> Residual Network (<b>ResNet</b>) <b>Loss:</b> Multi-task Loss (classification loss, bounding box regression loss and mask prediction loss) The proposed model was used to differentiate COVID-19 pneumonia from other pulmonary diseases and had the ability to classify objects into different classes, create bounding boxes, and generate a mask for the detected objects</p> <p><b>Comparison:</b> Decompose, Transfer &amp; Compose Residual Network-18 (<b>DeTraC ResNet-18</b>), Single Shot Detector (<b>SSD</b>), Corona Virus Disease X-Net (<b>COVIDX-Net</b>), Visual Geometry Group Network (<b>VGG-19</b>), Corona Virus Disease Network (<b>COVID-Net</b>), ResNet-50+Support Vector Machine (<b>SVM</b>) and Shallow CNN</p> <p><b>Datasets:</b> Cell images of one distinct organ (Lung) were taken from COVID-19 Dataset (Cohen et al. 2020)</p> <p><b>Parameters:</b> Accuracy, Specificity (SP), Recall, Precision and F1-score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby making it robust and capable enough to detect lungs with no false alarm as it offered both bounding box and instance segmentation. Further, the proposed model might be altered after noticing other abnormalities in the lungs. The accuracy and specificity measured for the COVID-19 Dataset were 96.98% and 97.36%, respectively</p> <p><b>Limitations:</b> The model's drawback is that more extensive datasets are needed for perfect detection</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Binary cross-entropy with logits Adam's optimizer was employed. The proposed model substituted the convolutional blocks of the traditional U-Net with the Multi-Level Dilated Residual (MLDR) blocks. Furthermore, the skip connection mechanism was altered by familiarizing the multi-level residual (MLR) network</p> <p><b>Comparison:</b> U-Net, U-Net++, Residual U-Net, ResDU-Net and MultiResU-Net</p> <p><b>Datasets:</b> Cell images of seven distinct organs (Colon, Brain) were taken from ISIC-2018 Dataset (Codella et al. 2019; Tschandl et al. 2018), ISBI-2012 Dataset (Arganda-Carreras et al. 2015; Cardona et al. 2010), MRI Dataset (Buda 2020; Buda et al. 2019), Glas-2015 Dataset (Ultrasound nerve segmentation 2016), and DSB-2018 Dataset (Caicedo et al. 2019; Data science bowl 2018)</p> <p><b>Parameters:</b> Dice coefficient (<b>DC</b>), Intersection over Union (<b>IoU</b>) and Hausdorff Distance (<b>HD</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby outperforming the approach by relative improvements of 1%, 1%, 1%, 4%, and 4%, respectively. Furthermore, saliency maps project that the proposed method focuses much better on the ROIs in images with complex backgrounds. The measured values of (DC = 0.98 ± 1.386, IoU = 0.93 ± 1.252, HD = 9.254 ± 1.75) for ISBI-2012 Dataset, (DC = 0.94 ± 0.0042, IoU = 0.91 ± 0.036, HD = 7.39 ± 0.064) for ISIC-2018 Dataset, (DC = 0.89 ± 0.005, IoU = 0.81 ± 0.0002, HD = 13.02 ± 0.0083) for MRI Dataset, (DC = 0.87 ± 0.0003, IoU = 0.80 ± 1.294, HD = 15.408 ± 0.0574) for Glas Dataset and (DC = 0.95 ± 0.0003, IoU = 0.91 ± 0.328, HD = 4.078 ± 0.0020) for DSB-2018 Dataset, respectively</p> <p><b>Limitations:</b> The segmentation is a challenging task for the model when the MRI images have a complex shape and structure</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Cross-entropy Adam's optimizer was employed. The proposed model comprised a contractile route to capture background and an asymmetric expansion path that offered specific localization and provided end-to-end training from among a small number of images</p> <p><b>Comparison:</b> Bidirectional Recurrent Neural Networks (<b>HA-BIRNN</b>), Support Vector Machine (<b>SVM</b>), Naive Bayes and Region Based Convolutional Neural Network (<b>R-CNN</b>)</p> <p><b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from The Cancer Genome Atlas (<b>TCGA</b>) Dataset (Tomczak et al. 2015); The Cancer Genome Atlas (<b>TCGA</b>) 2016)</p> <p><b>Parameters:</b> Accuracy, Precision, Recall and F1-Score</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby attaining a high percentage of IoU and accuracy with low losses. For the TCGA Dataset, the measured values were: precision = 92.00, accuracy = 92.50, and recall = 90.00</p> <p><b>Limitations:</b> Due to the skip connection in the model architecture, the proposed model does not perform well</p>
<b>Year: 2021</b> Gudhe et al. (2021) proposed a Multilevel Dilated Residual Neural Network ( <b>MILDNet</b> ) for the biomedical image segmentation task	<p><b>Year: 2021</b> Tarighat (2021) proposed a deep convolutional neural network by U-net for Breast Tumour Segmentation</p>

Table 1 (continued)

Paper details	
<b>Year: 2021</b> Zunair and Hamza (2021) proposed Sharp U-Net for binary and multi-class biomedical image segmentation	<p><b>Backbone:</b> U-Net <b>Loss:</b> Categorical Cross-entropy Adam optimizer was employed, and 7.8 million learnable parameters were used. To reduce feature mismatch, the proposed model fused decoder and encoder features via depth-wise convolution with a sharpening spatial kernel</p> <p>U-Net, Wide U-Net, TerausNet-16 and U-Net + ResNet-50</p> <p>Cell images of two distinct organs (Lung and Colon) were taken from Lung Segmentation Dataset (Li et al. 2020), Data Science Bowl 2018 Dataset (Caicedo et al. 2019), Data science bowl 2018), ISIC-2018 Dataset (Codella et al. 2018), COVID-19 CT Segmentation Dataset (Codella et al. 2018), ISBI-2012 Dataset (Arganda-Carreiras et al. 2015; Cardona et al. 2010) and CVC-ClinicDB Dataset (Bernal et al. 2015)</p> <p>Jaccard and Dice</p> <p>The proposed model was better in terms of performance as compared to other models, thereby enhancing artefacts throughout the network layers during the early stages of training. Additionally, other loss functions, such as the focal Tversky and distance-based losses, were used for training. The measured values of Jaccard=83.98±0.27, Dice=90.05±0.29) for CVC-ClinicDB Dataset, (Jaccard=91.21±0.67, Dice=93.52±0.91) for ISBI-2012 Dataset, (Jaccard=91.22±1.20, Dice=94.65±0.69) for COVID-19 CT Segmentation Dataset, (Jaccard=79.78±0.55, Dice=87.01±0.90) for ISIC-2018 Dataset, (Jaccard=89.60±0.15, Dice=95.40±0.10) for Data Science Bowl Dataset and (Jaccard=97.44±0.24, Dice=98.73±0.16) for Lung Segmentation Dataset, respectively</p> <p>Compared to U-Net, Sharp U-Net predicts segmented ROIs that aren't as well divided and aren't as good, but it makes a lot less noise and broken segmented ROIs</p>
<b>Year: 2021</b> Liu et al. (2021a) proposed Multiscale Connected Segmentation Network with Distance map and Contour information (MDC-Net) for nucleus segmentation in histopathology images	<p><b>Backbone:</b> ResNet-101 <b>Loss:</b> Cross-entropy</p> <p>The proposed method employed multiple short residual connections and dilated convolution with diverse dilation ratios to fuse feature maps from different scales for improved utilization of the context information and to enhance the receptive fields, respectively</p> <p>Mask Regional Convolutional Neural Network (Mask R-CNN), U-Net, DIST and Refine-Net</p> <p>Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from DATA ORGANS (Naylor et al. 2017) and DATA BREAST (Naylor et al. 2018)</p> <p>F1 Score, Aggregated Jaccard Index (AJI) and Hausdorff distance</p> <p>The proposed model was better in terms of performance as compared to other models, thereby enhancing the morphology and boundary information of a nucleus to successfully alleviate "touching and partial overlap" and acquire perfect nucleus segmentation. For the DATA ORGANS dataset, the measured values of Aji, Hausdorff, and F1 distance were 0.5951, 6.7742, and 0.8234, respectively, and for the DATA BREAST dataset, they were 0.6103, 3.3300, and 0.8457, respectively</p> <p>The biggest problem is that the smallest, most tightly packed nuclei are the hardest to separate</p>
<b>Year: 2021</b> Aatresh et al. (2021a) Proposed a LiverNet for automatic diagnosis of sub-types of liver hepatocellular carcinoma cancer from H&E-stained liver histopathology images	<p><b>Backbone:</b> Not mentioned <b>Loss:</b> Categorical cross-entropy</p> <p>The proposed model used two different convolution layers for feature extraction before initiating the preliminary max-pool operation and further employed CBAM blocks and residual blocks for resourceful extraction of information. Additionally, hyper-column techniques coupled with ASPP blocks guarantee multi-scale feature extraction and information retrieval for supplementary processing</p> <p>Residual Network (ResNet), DenseNet, Inception ResNetV2, Inception Recurrent Residual Convolutional Neural Network (IR-RCNN) and BreastNet. The model has 573,936 parameters</p> <p>Cell images of one distinct organ (Liver) was taken from TCGA Liver dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016), Kasturba Medical College (KMC) Dataset (Kasturba Medical College 2021)</p> <p>F1-Score, IoU (JI), Accuracy, Precision and Recall</p> <p>The proposed model was better in terms of performance as compared to other models, making it capable of handling multi-class HCC histopathology image classification among various sub-types of liver HCC tumor. Moreover, the dataset used in the proposed work was restricted in terms of size and variance, making it a limitation of the proposed module. The measured values of precision of 0.9093, recall of 0.9093, F1-score of 0.9093, and IoU of 0.8360 on the KMC dataset and TCGA-Liver dataset had an accuracy of 97.72%, precision of 0.9772, recall of 0.9772, F1-score of 0.9772, and IoU of 0.9561</p> <p>The dataset being constrained in size and variation is one of the study's limitations. This is mostly because there aren't many liver histopathology imaging data sets that are available to the public</p>



Table 1 (continued)

Paper details	
<b>Year:</b> 2021	<b>Backbone:</b> Not mentioned
Aatresh et al. (2021b) proposed <b>Kidney-SegNet</b> for nuclei segmentation of histopathology images	<b>Loss:</b> Global loss
	The proposed model considered U-net style encoder-decoder architecture with an attention gating (AG) mechanism and used 1,359,652 parameters. All standard convolutions with DICE units were substituted with the intention of making the complete pipeline more competent. Skip connections with attention gates amid the encoder and its corresponding decoder layers were announced to counter the loss of information owing to multiple max-pool operations
	<b>Comparison:</b> CNN Model with One Convolutional Layer (CNN-1), Segmentation Network (SegNet), U-Net, CUMed Vision, Attention U-Net, DIST, DeepLab V3+, U-Net + +, Separable Convolutional Pyramid Pooling Network (SCPP-Net) and Arous Spatial Pyramid Pooling U-Net (ASPP U-Net)
	<b>Datasets:</b> Cell images of two distinct organs (Breast, Kidney) were taken from H&E-Stained Kidney Dataset (Irshad et al. 2015) and Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)
	<b>Parameters:</b> F1 score and Aggregated Jaccard Index (AJI)
	<b>Inference:</b> The proposed model was better in terms of performance as compared to other models, thereby revealing that the computational complexity and memory requirement of the proposed architecture were very effective. The measured values are from H&E-stained histopathology images, with the kidney dataset having an f1 score and an AJI of 0.9294 and 0.8688, respectively, and the breast dataset having an f1 score and an AJI of 0.8243 and 0.7039, respectively
	<b>Limitations:</b> For the proposed model, segmentation is a challenging task when the images have a complex shape and structure
	<b>Features:</b> <b>Backbone:</b> Not mentioned
	<b>Loss:</b> Binary Cross-entropy (BCE)
	Adam optimizer was employed, and the proposed model used residual connections between bottleneck layers
	<b>Comparison:</b> Horizontal and Vertical Distance Network (HoVer-Net), Fully Convolutional Network-8 (FCN-8) + WS, U-Net, Mask Regional Convolutional Neural Network (Mask-RCNN), Deep Convolutional Auto-encoder Network (DCAN), Micro-Net, DIST, SegNet, Pix2Pix, Graph Convolutional Network (GCN), Multi-Scale Residual Convolutional Neural Network (MRCNN) and Deep Panoptic
	<b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Triple Negative Breast Cancer (TNBC) dataset (Naylor et al. 2018)
	<b>Parameters:</b> Area Under the Curve (AUC) and Intersection over Union (IOU)
	<b>Inference:</b> The proposed model was better in terms of performance as compared to other models. The MobileNetV2-based U-net model provided an accuracy value of 0.9460, an AUC value of 0.9179, and an IOU value of 0.5141. The MobileNetV2-based U-net model with data augmentation gave an accuracy value of 0.9731, an AUC value of 0.9821, and an IOU value of 0.5931
	<b>Limitations:</b> When the model is trained without the augmented dataset, it doesn't do a very good job of separating things
	<b>Features:</b> <b>Backbone:</b> ResNet-50
	<b>Loss:</b> Adversarial Segmentation Loss and Global loss
	Adam's optimizer was employed. The proposed model used AdaIN as a generator, along with instance deformation and style adaptation, to create nuclei images with advanced variations and realism, resulting in significantly improved nuclear pleomorphisms and texture patterns
	<b>Comparison:</b> (i) Evaluated on the TCGA Dataset - Kumar et al. (2017), DIST, Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN), Cell Region Based CNN v2 (Cell R-CNN v2) and Cell Region Based CNN v3 (Cell R-CNN v3)
	(ii) Evaluated on the Cell17 Dataset - Pix2Pix, Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN) and Liu et al. (2019)
	(iii) Evaluated on the TNBC Dataset - Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN), Cell Region Based CNN v2 (Cell R-CNN v2) and Cell Region Based CNN v3 (Cell R-CNN v3)
	<b>Datasets:</b> Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016), Cell17 Dataset (Vu et al. 2018) and Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)
	<b>Parameters:</b> Aggregated Jaccard Index (AJI), Pixel-F1, Obj-F1 and Dice
	<b>Inference:</b> The proposed model outperformed other models in terms of performance, indicating that the synthetic images aided in improving instance segmentation performance. The measured values of (AJI = 0.6346 ± 0.0674, Pixel-F1 = 0.8180 ± 0.0131, Obj-F1 = 0.8379 ± 0.0292) for TCGA Dataset, (Pixel-F1 = 0.8622 ± 0.0087, Dice = 0.8216 ± 0.0103) for Cell17 Dataset, (AJI = 0.6316 ± 0.0597, Pixel-F1 = 0.8231 ± 0.0137) for TNBC Dataset, respectively
	<b>Limitations:</b> When the images have a complex shape and structure, segmentation becomes a difficult task for the model
<b>Year:</b> 2021	<b>Backbone:</b> Not mentioned
Kanadath et al. (2021) proposed a <b>MobileNetV2</b> for the segmentation of nuclei regions from Triple Negative Breast Cancer (TNBC) histopathology images	<b>Loss:</b> Binary Cross-entropy (BCE)
	Adam optimizer was employed, and the proposed model used residual connections between bottleneck layers
	<b>Comparison:</b> Horizontal and Vertical Distance Network (HoVer-Net), Fully Convolutional Network-8 (FCN-8) + WS, U-Net, Mask Regional Convolutional Neural Network (Mask-RCNN), Deep Convolutional Auto-encoder Network (DCAN), Micro-Net, DIST, SegNet, Pix2Pix, Graph Convolutional Network (GCN), Multi-Scale Residual Convolutional Neural Network (MRCNN) and Deep Panoptic
	<b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Triple Negative Breast Cancer (TNBC) dataset (Naylor et al. 2018)
	<b>Parameters:</b> Area Under the Curve (AUC) and Intersection over Union (IOU)
	<b>Inference:</b> The proposed model was better in terms of performance as compared to other models. The MobileNetV2-based U-net model provided an accuracy value of 0.9460, an AUC value of 0.9179, and an IOU value of 0.5141. The MobileNetV2-based U-net model with data augmentation gave an accuracy value of 0.9731, an AUC value of 0.9821, and an IOU value of 0.5931
	<b>Limitations:</b> When the model is trained without the augmented dataset, it doesn't do a very good job of separating things
	<b>Features:</b> <b>Backbone:</b> ResNet-50
	<b>Loss:</b> Adversarial Segmentation Loss and Global loss
	Adam's optimizer was employed. The proposed model used AdaIN as a generator, along with instance deformation and style adaptation, to create nuclei images with advanced variations and realism, resulting in significantly improved nuclear pleomorphisms and texture patterns
	<b>Comparison:</b> (i) Evaluated on the TCGA Dataset - Kumar et al. (2017), DIST, Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN), Cell Region Based CNN v2 (Cell R-CNN v2) and Cell Region Based CNN v3 (Cell R-CNN v3)
	(ii) Evaluated on the Cell17 Dataset - Pix2Pix, Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN) and Liu et al. (2019)
	(iii) Evaluated on the TNBC Dataset - Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN), Cell Region Based CNN v2 (Cell R-CNN v2) and Cell Region Based CNN v3 (Cell R-CNN v3)
	<b>Datasets:</b> Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016), Cell17 Dataset (Vu et al. 2018) and Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)
	<b>Parameters:</b> Aggregated Jaccard Index (AJI), Pixel-F1, Obj-F1 and Dice
	<b>Inference:</b> The proposed model outperformed other models in terms of performance, indicating that the synthetic images aided in improving instance segmentation performance. The measured values of (AJI = 0.6346 ± 0.0674, Pixel-F1 = 0.8180 ± 0.0131, Obj-F1 = 0.8379 ± 0.0292) for TCGA Dataset, (Pixel-F1 = 0.8622 ± 0.0087, Dice = 0.8216 ± 0.0103) for Cell17 Dataset, (AJI = 0.6316 ± 0.0597, Pixel-F1 = 0.8231 ± 0.0137) for TNBC Dataset, respectively
	<b>Limitations:</b> When the images have a complex shape and structure, segmentation becomes a difficult task for the model
<b>Year:</b> 2021	<b>Backbone:</b> Not mentioned
Gong et al. (2021) proposed <b>Style-Consistent Generation Method</b> for nuclei instance segmentation in histology images	<b>Loss:</b> Binary Cross-entropy (BCE)
	Adam optimizer was employed, and the proposed model used residual connections between bottleneck layers
	<b>Comparison:</b> Horizontal and Vertical Distance Network (HoVer-Net), Fully Convolutional Network-8 (FCN-8) + WS, U-Net, Mask Regional Convolutional Neural Network (Mask-RCNN), Deep Convolutional Auto-encoder Network (DCAN), Micro-Net, DIST, SegNet, Pix2Pix, Graph Convolutional Network (GCN), Multi-Scale Residual Convolutional Neural Network (MRCNN) and Deep Panoptic
	<b>Datasets:</b> Cell images of one distinct organ (Breast) were taken from Triple Negative Breast Cancer (TNBC) dataset (Naylor et al. 2018)
	<b>Parameters:</b> Area Under the Curve (AUC) and Intersection over Union (IOU)
	<b>Inference:</b> The proposed model was better in terms of performance as compared to other models. The MobileNetV2-based U-net model provided an accuracy value of 0.9460, an AUC value of 0.9179, and an IOU value of 0.5141. The MobileNetV2-based U-net model with data augmentation gave an accuracy value of 0.9731, an AUC value of 0.9821, and an IOU value of 0.5931
	<b>Limitations:</b> When the model is trained without the augmented dataset, it doesn't do a very good job of separating things
	<b>Features:</b> <b>Backbone:</b> ResNet-50
	<b>Loss:</b> Adversarial Segmentation Loss and Global loss
	Adam's optimizer was employed. The proposed model used AdaIN as a generator, along with instance deformation and style adaptation, to create nuclei images with advanced variations and realism, resulting in significantly improved nuclear pleomorphisms and texture patterns
	<b>Comparison:</b> (i) Evaluated on the TCGA Dataset - Kumar et al. (2017), DIST, Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN), Cell Region Based CNN v2 (Cell R-CNN v2) and Cell Region Based CNN v3 (Cell R-CNN v3)
	(ii) Evaluated on the Cell17 Dataset - Pix2Pix, Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN) and Liu et al. (2019)
	(iii) Evaluated on the TNBC Dataset - Mask Regional Convolutional Neural Network (Mask R-CNN), Cell Region Based CNN (Cell R-CNN), Cell Region Based CNN v2 (Cell R-CNN v2) and Cell Region Based CNN v3 (Cell R-CNN v3)
	<b>Datasets:</b> Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from The Cancer Genome Atlas (TCGA) Dataset (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016), Cell17 Dataset (Vu et al. 2018) and Triple Negative Breast Cancer (TNBC) Dataset (Naylor et al. 2018)
	<b>Parameters:</b> Aggregated Jaccard Index (AJI), Pixel-F1, Obj-F1 and Dice
	<b>Inference:</b> The proposed model outperformed other models in terms of performance, indicating that the synthetic images aided in improving instance segmentation performance. The measured values of (AJI = 0.6346 ± 0.0674, Pixel-F1 = 0.8180 ± 0.0131, Obj-F1 = 0.8379 ± 0.0292) for TCGA Dataset, (Pixel-F1 = 0.8622 ± 0.0087, Dice = 0.8216 ± 0.0103) for Cell17 Dataset, (AJI = 0.6316 ± 0.0597, Pixel-F1 = 0.8231 ± 0.0137) for TNBC Dataset, respectively
	<b>Limitations:</b> When the images have a complex shape and structure, segmentation becomes a difficult task for the model

Table 1 (continued)

Year: 2021	Paper details
Liu et al. (2021b)	proposed a melanocytic proliferation segmentation pipeline framework for Melanocytic Proliferation Segmentation in Histopathology Images
<b>Features:</b>	<p>Backbone: ResNet-101 + Feature Pyramid Network (FPN)            Loss: Cross-entropy (CE), Weighted Cross-entropy (WCE), and Focal Loss (FL)            Stochastic Gradient Descent (SGD) was used as an optimizer. The proposed model adopts a framework based on Mask Regional Convolutional Neural Network (Mask R-CNN) and consists of two major components: data annotation, pre-processing procedures, and a melanocytic proliferation segmentation model with patch stitching</p>
<b>Comparison:</b>	State-of-the-Art (SOTA) Autoencoder
<b>Datasets:</b>	Their dataset consists of 227 regions of interest (ROI) images extracted from haematoxylin and eosin (H&E) stained slides of skin biopsy images at 10 X magnification and diagnosed by three expert pathologists, who all agreed on the consensus diagnosed and selected the ROIs
<b>Parameters:</b>	Dice, mean Intersection Over Union (mIOU), Accuracy, Sensitivity and Specificity
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models; thus, it required partially labeled datasets, which vastly reduced the data annotation cost, further paving the way towards assisting in diagnosis by pathologists and aiding in downstream computer vision analysis. The measured values of the MS COCO Dataset are: Dice = 0.719, mIOU = 0.740, Accuracy = 0.927, Sensitivity = 0.751, and Specificity = 0.952
<b>Limitations:</b>	Due to the complexity and expense of gathering and annotating training data, they are constrained by the lack of such data. In the sample with limited labeling, the scarcity is significantly worse. In actuality, there are only 130 images in the sample that contain annotated melanocytic proliferations. Mask R-CNN typically fails to recognize objects with low-resolution motion blur, like hands
<b>Features:</b>	<p><b>Backbone:</b> Not mentioned  <b>Loss:</b> Binary Cross-entropy (BCE)            The Adam optimizer was used with a model that had 280 K trainable parameters. They used their workflow to detect tumor-infiltrating lymphocytes in images of breast and colorectal cancer tissue stained with haematoxylin and eosin (H&amp;E). For the lymphocyte classification task, they utilized traditional machine learning approaches: a random forest classifier, a multilayer perceptron, and a CNN. The modified Micro-Net model architecture incorporated texture convolutional blocks, which facilitated relevant feature extraction for the autoencoder</p>
<b>Comparison:</b>	Janowczyk and Madabhushi (2016), Alom et al. (2019)
<b>Datasets:</b>	Cell images of two distinct organs (Breast and Colon) are taken from CRC Histo Phenotypes (CRCHP) Dataset (Janowczyk and Madabhushi 2016) and Breast Cancer Dataset (JAN) (Janowczyk and Madabhushi 2016)
<b>Parameters:</b>	Accuracy, Precision, Recall and F1-Score
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thus displaying decent generalization properties, eradicating overfitting, and being able to be employed in multi-class nuclei segmentation and identification tasks. The measured values of the Same Dataset are: accuracy = 0.71, precision = 0.76, recall = 0.75, and F1-Score = 0.70
<b>Limitations:</b>	When the images have a complex shape and structure, segmentation becomes a difficult task for the model
<b>Features:</b>	<p><b>Backbone:</b> Not mentioned  <b>Loss:</b> Mean Squared Error (MSE), Structural Similarity (SSIM) index and Mean Absolute Error (MAE)            Adam's optimizer was employed. The proposed model consisted of a framework for image reconstruction with a customized reconstruction loss function, followed by a module to classify each image patch as a tumour versus a non-tumour</p>
<b>Comparison:</b>	Residual Network-101 (ResNet-101), Visual Geometric Group-19 (VGG-19), Densely Connected Convolutional Networks (DenseNet), Inception V3, U-Net, Segmentation Network (SegNet), DeepLab v3, Pyramid Scene Parsing Network (PSP-Net), Refine-Net and Mobile-UNet
<b>Datasets:</b>	Cell images of one distinct organ (Liver) was taken from the dataset were comes from PAIP challenge 2019 that is a part of MICCAI 2019 grand challenges (PAIP2019 2019). It included 50 fully annotated H&E WSIs of liver tissue at 20X magnification
<b>Parameters:</b>	Accuracy, Dice Similarity, Precision, Recall and F1-Score
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models, thus proving its effectiveness for viable tumour area segmentation. The loss function of MSE + SSIM produced an accuracy of 0.92, whereas the loss function of MSE + SSIM + MAE produced an accuracy of 0.95. Similarly, the WSI tumour segmentation performance assessed by dice value improved from 0.56 for SSIM loss to 0.76 for SSIM + MAE-based loss, and finally to 0.87 for MSE + SSIM + MAE-based loss
<b>Limitations:</b>	The accuracy of the classification is then diminished by the poor reconstruction outcome, and HistoCAE1 can only process data at a single 10 X magnification
<b>Year: 2021</b>	
Roy et al. (2021)	proposed Histopathology Convolutional Autoencoder (HistoCAE) for segmentation of viable tumour regions in liver

Table 1 (continued)

Paper details	
<p><b>Year: 2021</b> Dogan et al. (2021) proposed two-phase approach for high-accuracy automatic pancreas segmentation in Computed Tomography (CT) imaging</p>	<p><b>Backbone:</b> Not mentioned <b>Loss:</b> Binary Cross-entropy (BCE) The proposed approach contained two important phases: Pancreas localization, where the Mask regional convolutional neural network (<b>Mask R-CNN</b>) model was executed by observing the rough pancreas position on the 2D CT slice; pancreas segmentation, wherein the candidate pancreas was refined using 3D U-Net on the 2D sub-CT slices generated in the first phase</p> <p><b>Comparison:</b> Hierarchical Coarse-to-Fine, ConvNets, Holistically Nested Network, Coarse-to-Fine, Recurrent Neural Contextual Learning, Saliency Transformation Network, Bayesian Model, 3D Coarse-to-Fine, 3D U-Net, Projective Adversarial Network, Model Driven Stack based FCN, Cascaded 3D FCN, Ensemble-based FCN and 2D U-Net</p> <p><b>Datasets:</b> Cell images of three distinct organs (Brain, Liver and Kidney) were taken from NIH-82 Dataset (Roth et al. 2018; Zhou et al. 2016)</p> <p><b>Parameters:</b> Dice Similarity Coefficient (<b>DSC</b>), Jaccard Index (<b>JI</b>), Precision (<b>PRE</b>), Recall (<b>REC</b>), Pixel Accuracy (<b>ACC</b>), Specificity (<b>SPE</b>), Receiver Operating Characteristics (<b>ROC</b>) and Area under curve (<b>AUC</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, wherein models with lesser GPU capacity turned out to be capable of achieving more fruitful outcomes as it employed two phases, thereby drastically reducing the processing memory space. The measured values of the NIH-82 Dataset are: DSC=86.15%, JI=75.93%, PRE=86.23%, REC=86.27%, and ACC=99.95%</p> <p><b>Limitations:</b> The capacity of the GPU is rather tiny in comparison to other investigations done for pancreatic segmentation. For pancreas segmentation, the suggested approach's performance assessment methodologies yield poorer findings than those of the previous research</p> <p><b>Backbone:</b> Not mentioned <b>Loss:</b> Weighted Cross-entropy (WCE) Adam's optimizer was employed. The proposed pre-processing pipeline was applied to HSI data prior to the tissue segmentation step, comprising 5 main operations involving outlier removal, background identification, spectral smoothing, data normalization, dimensional reduction, and spatial smoothing</p> <p><b>Comparison:</b> Inception-v4 CNN, Multitask U-Net framework, SVM with polynomial kernel (Schoels et al. 2014), SVM with polynomial kernel (Schoels et al. 2017) and SVM with RBF kernel (Maktabi et al. 2020)</p> <p><b>Datasets:</b> Cell images of two distinct organs (Brain and Liver) were taken from HSI Dataset (Maktabi et al. 2020)</p> <p><b>Parameters:</b> Mean Accuracy and F1-Score</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance, but it was limited by some extrinsic factors, such as unintentional mislabeling in the ground truth data. The measured HSI Dataset, Mean Accuracy, and F1-Score values for the bile duct were 0.901 and 0.787, respectively, and 0.696 and 0.476 for the parathyroid</p> <p><b>Limitations:</b> The unintentional misclassification of the ground truth data is just one of several external variables that contribute to the suggested method's drawbacks. The problems were caused by things on the inside, like the fact that the tissue composition was different in each patient</p>
<p><b>Year: 2021</b> Cervantes-Sanchez et al. (2021) proposed a pre-processing pipeline for automatic tissue segmentation of hyperspectral images in liver and head neck surgeries</p>	<p><b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned The proposed method encompassed three foremost phases: pre-processing, brain tumour segmentation via k-means clustering, and classification of tumours into categories (benign/malignant) by means of MRI data through a fine-tuned VGG19 (i.e., 19-layered Visual Geometric Group) model</p> <p><b>Comparison:</b> Novel 2D CNN, Seetha and Raja (2018), Novel 3D CNN, Pre-trained VGG-19, Özyurt et al. (2019), Pre-trained inception V3 and 3D CNN</p> <p><b>Datasets:</b> Cell images of one distinct organ (Brain) were taken from BraTS 2015 Benchmark Dataset (Menze et al. 2014)</p> <p><b>Parameters:</b> Accuracy</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance, and it was easily trained using a synthetic data augmentation procedure, making it a trustworthy model for assisting radiologists and medical experts. The measured values of accuracy were 90.03% before data augmentation and 94.06% after synthetic data augmentation for the BraTS 2015 benchmark dataset</p> <p><b>Limitations:</b> When the shape and structure of the images are complicated, it is hard for the model to separate them</p>
<p><b>Year: 2021</b> Khan et al. (2021) proposed a deep learning approach to classify brain tumours using an MRI data analysis to assist practitioners</p>	<p><b>Backbone:</b> Not mentioned <b>Loss:</b> Not mentioned The proposed method encompassed three foremost phases: pre-processing, brain tumour segmentation via k-means clustering, and classification of tumours into categories (benign/malignant) by means of MRI data through a fine-tuned VGG19 (i.e., 19-layered Visual Geometric Group) model</p> <p><b>Comparison:</b> Novel 2D CNN, Seetha and Raja (2018), Novel 3D CNN, Pre-trained VGG-19, Özyurt et al. (2019), Pre-trained inception V3 and 3D CNN</p> <p><b>Datasets:</b> Cell images of one distinct organ (Brain) were taken from BraTS 2015 Benchmark Dataset (Menze et al. 2014)</p> <p><b>Parameters:</b> Accuracy</p> <p><b>Inference:</b> The proposed model outperformed other models in terms of performance, and it was easily trained using a synthetic data augmentation procedure, making it a trustworthy model for assisting radiologists and medical experts. The measured values of accuracy were 90.03% before data augmentation and 94.06% after synthetic data augmentation for the BraTS 2015 benchmark dataset</p> <p><b>Limitations:</b> When the shape and structure of the images are complicated, it is hard for the model to separate them</p>

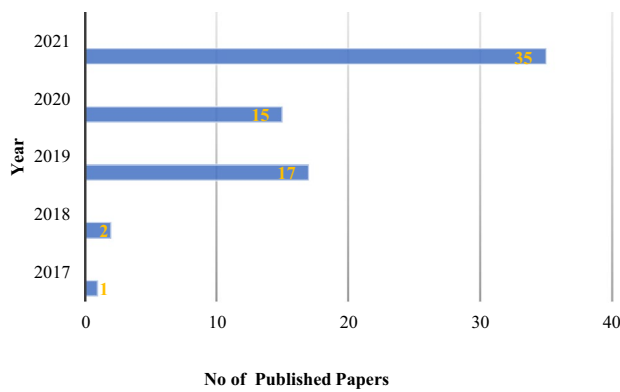
Table 1 (continued)

Paper details	
<p><b>Year: 2021</b> Dinh et al. (2021) proposed an <b>Efficient U-net++</b> for Nuclei Segmentation of Histopathology Images</p>	<p><b>Features:</b> <b>Backbone:</b> EfficientNet <b>Loss:</b> Dice Coefficient Loss function Adam's optimizer was employed. The proposed model used augmentations as the library for the training data augmentation pipeline to enrich their data, ranging from 19 to 5320 images for training and 5 images to 1400 images for validation. For optimization purposes, the squeeze-and-excitation components of MBCConv were explored Lagree et al. (2021)</p> <p><b>Comparison:</b> Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from Multi-Organ Nuclei Segmentation (MoNuSeg) (Kumar et al. 2017, 2020) and Triple Negative Breast Cancer (TNBC) (Naylor et al. 2018)</p> <p><b>Datasets:</b> Recall, F1-Score, Precision</p> <p><b>Parameters:</b> The proposed model was better in terms of performance as compared to other models, thereby achieving a large accuracy margin. The measured values are (Recall = 0.9272, F1-Score = 0.8008, and Precision = 0.7384) and (Recall = 0.9343, F1-Score = 0.6785, and Precision = 0.5931) for the MoNuSeg and TNBC Datasets, respectively</p> <p><b>Inference:</b> More training data fed to the network is required to improve accuracy. In order to increase generalization and avoid being overfit, the network needs more training data</p> <p><b>Limitations:</b></p>
<p><b>Year: 2021</b> Ke et al. (2021) proposed a <b>computer-aided Cytology Image Diagnostic System</b> of abnormalities in gynaecologic cytopathology</p>	<p><b>Features:</b> <b>Backbone:</b> ResNet-50 <b>Loss:</b> Cross-entropy Adam's optimizer was employed. The overall architecture of the proposed system entailed five important functional components: the segmentation model, the classification model, the spatial correlation model, the nuclear area correction model, and the aggregation model</p> <p><b>Comparison:</b> U-Net, Mask Regional Convolutional Neural Network (<b>Mask R-CNN</b>), U-Net++ +</p> <p><b>Datasets:</b> They privately collected and manually annotated dataset of 130 cytological whole-slide images from Shanxi Tumour Hospital (Kather et al. 2019)</p> <p><b>Parameters:</b> Pixel Accuracy, Mean Pixel Accuracy and Mean IoU</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, highlighting that the proposed methods were accomplished to effectively extract, interpret, and quantify morphological features. The measured values of Pixel Accuracy = 0.974 ± 0.001, Mean Pixel Accuracy = 0.955 ± 0.007 and Mean IoU = 0.913 ± 0.007</p> <p><b>Limitations:</b> Blood, folded cytoplasmic borders, or large clusters of overlapping epithelial cells do not reduce the sensitivity of the diagnosis. Cellular-level false positives are possible. Overlapping or fuzzy nuclei were sometimes misdiagnosed as dysplastic. Signal plane focusing made it hard to rule out inaccurate classifications on out-of-focus cells</p> <p><b>Features:</b> <b>Backbone:</b> Not mentioned <b>Loss:</b> Exponential Logarithmic Loss The Adam optimizer was used, and the model contained and used 20,306,691 parameters. The proposed model R<sup>2</sup>U3D architecture could efficiently process volumetric data and had been inspired by the concept of volumetric image segmentation</p> <p><b>Comparison:</b> V-Net and Extended V-Net</p> <p><b>Datasets:</b> Cell images of one distinct organ (Lung) are taken from Lung Nodule Analysis 2016 (<b>LUNA16</b>) (LUNA16—Home 2020) and VESSEL Segmentation in the Lung 2012 (<b>VESSEL12</b>) (VESSEL12—Home 2020)</p> <p><b>Parameters:</b> Soft Dice Similarity Coefficient (<b>Soft-DSC</b>)</p> <p><b>Inference:</b> The proposed model was better in terms of performance as compared to other models, highlighting that training the R<sup>2</sup>U3D model with a smaller number of CT scans, i.e., 100 scans, without applying data augmentation achieved an outstanding result. The measured values of Soft-DSC are 0.9920 and 0.9859 for the VESSEL12 and LUNA16 datasets, respectively</p> <p><b>Limitations:</b> Due to the skip connection in the model architecture, the proposed model does not perform well</p>

Table 1 (continued)

Paper details	
<b>Year:</b> 2021	<b>Backbone:</b> Not mentioned
Li et al. (2021b) proposed Bagging Ensemble Deep segmentation (BEDs) method for nuclei segmentation	<b>Loss:</b> Not mentioned The proposed method applies testing stage stain normalization to lessen the gap in presence between testing and training images. For the test images, patches of $256 \times 256$ pixels that overlapped by 20 pixels were tiled together. Further, after segmenting all patches, they were merged to the original image scale by performing a bitwise AND operation for the overlapped regions
<b>Comparison:</b>	Benchmark Model
<b>Datasets:</b>	Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from Multi-Organ Nuclei Segmentation (MoNuSeg) dataset (Kumar et al. 2017, 2020)
<b>Parameters:</b>	Dice Similarity Coefficients (DSC) and F1-Score
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models where aggregating self-ensemble learning and testing stage augmentation improved the robustness of nucleus segmentation. The measured values of pixel-wise DSC (mean = 0.8177 and median = 0.8192) and F1-Score = 0.8836 for the MoNuSeg dataset
<b>Limitations:</b>	The diverse image structure, even when using the same Hematoxylin & Eosin (H&E) staining, is a crucial uncertainty when segmenting nuclei on pathology images. When the pictures are complicated in shape and structure, segmentation becomes a difficult challenge for the model
<b>Features:</b>	<b>Backbone:</b> Horizontal and Vertical Distance Network (HoVer-Net) <b>Loss:</b> Bending Loss Bend-Net was proposed to address the challenges faced by the overlapped nuclei segmentation as it lacked clear boundaries, similar background textures, and large size and morphology variations. Further, bending loss was proposed that basically defined high penalties for concave contour points having large curvatures and small penalties for convex contour points with small curvatures, thereby avoiding contour generation that encompasses multiple nuclei
<b>Comparison:</b>	Fully Convolutional Neural Network (FCN-8), U-Net, Segmentation Network (SegNet), Deep Convolutional Auto-encoder Network (DCAN), DIST, Horizontal and Vertical Distance Network (HoVer-Net), Micro-Net and BEND
<b>Datasets:</b>	Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from Multi-Organ Nuclei Segmentation v1 (MoNuSegv1) (Kumar et al. 2017, 2020), CoNSEP (Lu et al. 2018)
<b>Parameters:</b>	Dice, Aggregated Jaccard Index (AJI), Recognition Quality (RQ), Segmentation Quality (SQ), Panoptic Quality (PQ)
<b>Inference:</b>	The proposed model outperformed other models in terms of performance, with experimental results confirming that the bending loss effectively progressed, thereby improving the overall performance of nuclei segmentation. The measured values of AJI = 0.578, Dice = 0.851, RQ = 0.709, SQ = 0.781, and PQ = 0.555 for CoNSEP and AJI = 0.635, Dice = 0.832, RQ = 0.780, SQ = 0.771, and PQ = 0.601 for the MoNuSegv1 dataset, respectively
<b>Limitations:</b>	The problem is that it was hardest to separate the smallest and most tightly packed nuclei
<b>Features:</b>	<b>Backbone:</b> U-Net <b>Loss:</b> Binary Cross-Entropy (BCE) Adam optimizer was employed and the model had 8.7 M parameters. Proposed architecture comprised of high-resolution encoder ASPP bottleneck path for multi-level feature extraction avoiding internal covariate shift
<b>Comparison:</b>	U-Net, SegNet, Attention U-Net, DIST and Atrous Spatial Pyramid Pooling U-Net (ASPP U-Net) (2020)
<b>Datasets:</b>	Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from kidney dataset (Irshad et al. 2015), Triple Negative Breast Cancer (TNBC) dataset (Naylor et al. 2018), and MoNuSeg dataset (Kumar et al. 2017, 2020)
<b>Parameters:</b>	F1-Score and Aggregated Jaccard Index (AJI)
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models wherein residual learning as well as encoder-decoder architecture were effectively influenced by integrating wide and deep network paths that reinforced the intermediate features. The measured values were (F1-Score = 0.9684, AJI = 0.9394), (F1-Score = 0.8419, AJI = 0.7282), and (F1-Score = 0.8344, AJI = 0.7169) for the Kidney, TNBC, and Multiple Organs Multi-Disease Histopathology Datasets, respectively
<b>Limitations:</b>	The proposed model produced good results, but some histopathological images struggle with segmenting overlapping nuclei, and the published findings are insufficient for therapeutic use. Some nuclei have disappeared or have unclear borders. Complex histopathological images can result in over- and under-segmented nuclei
<b>Year:</b> 2021	<b>Backbone:</b> U-Net
Chanchal et al. (2021a) proposed CNN-based architecture called Atrous Spatial Pyramid Pooling U-Net (ASPP U-Net) for nuclei segmentation of histopathology images	<b>Loss:</b> Binary Cross-Entropy (BCE) Adam optimizer was employed and the model had 8.7 M parameters. Proposed architecture comprised of high-resolution encoder ASPP bottleneck path for multi-level feature extraction avoiding internal covariate shift
<b>Comparison:</b>	U-Net, SegNet, Attention U-Net, DIST and Atrous Spatial Pyramid Pooling U-Net (ASPP U-Net) (2020)
<b>Datasets:</b>	Cell images of seven distinct organs (Breast, Kidney, Liver, Prostate, Bladder, Stomach and Colorectal) were taken from kidney dataset (Irshad et al. 2015), Triple Negative Breast Cancer (TNBC) dataset (Naylor et al. 2018), and MoNuSeg dataset (Kumar et al. 2017, 2020)
<b>Parameters:</b>	F1-Score and Aggregated Jaccard Index (AJI)
<b>Inference:</b>	The proposed model was better in terms of performance as compared to other models wherein residual learning as well as encoder-decoder architecture were effectively influenced by integrating wide and deep network paths that reinforced the intermediate features. The measured values were (F1-Score = 0.9684, AJI = 0.9394), (F1-Score = 0.8419, AJI = 0.7282), and (F1-Score = 0.8344, AJI = 0.7169) for the Kidney, TNBC, and Multiple Organs Multi-Disease Histopathology Datasets, respectively
<b>Limitations:</b>	The proposed model produced good results, but some histopathological images struggle with segmenting overlapping nuclei, and the published findings are insufficient for therapeutic use. Some nuclei have disappeared or have unclear borders. Complex histopathological images can result in over- and under-segmented nuclei





**Fig. 1** Year-wise publishes papers of Nucleus Segmentation

The year wise number of published papers over Nucleus Segmentation has been presented in Fig. 1, which clearly demonstrates that Nucleus Segmentation is paving its way in the field of research.

## 4.2 Analysis based on dataset

This sub-section highlights a brief description about some of the extensively used Nucleus Segmentation datasets encountered while performing the literature survey as depicted in Table 1 namely TCGA (Tomczak et al. 2015; The Cancer Genome Atlas (TCGA) 2016), TNBC (Naylor et al. 2018), Herlev dataset (Jantzen et al. 2005), MS COCO (Lin et al. 2015), MoNuSeg (Kumar et al. 2017, 2020), DSB2018 (Caicedo et al. 2019; Data science bowl 2018), KMC Liver dataset (Kasturba Medical College 2021), and PanNuke dataset (Gamper et al. 2019, 2020), respectively.

### (i) The cancer genome atlas (TCGA) dataset:

The TCGA dataset is sponsored project wherein the researcher aims to analyse and produces an atlas of cancer genomic profiles (openly available datasets (The Cancer Genome Atlas (TCGA) 2016)) with over 20,000 cases of 33 types of cancer acknowledged till date. Kumar et al. for the purpose of nuclear segmentation task (Kumar et al. 2017) generated ground truths by picking around 44 WSIs of multiple organs especially images collected from seven different organs, including the bladder, breast, colon, kidney, liver, prostate, and stomach.

### (b) Triple negative breast cancer (TNBC) dataset:

For breast cancer histopathology images, Naylor et al. presented this dataset that deals with the type of breast cancer in which the cancer cells do not have oestrogen or progesterone receptors and yield in adequate amounts of the protein called HER2 and presented nuclear segmentation technique (Naylor et al. 2018) for the same. TNBC

encompasses 50 H&E-stained images with  $512 \times 512$  resolution and 4022 annotated nuclei. Entire images of TNBC are extracted from 11 triple negative breast cancer patients, and comprises of several cell types such as myoepithelial breast cells, endothelial cells and inflammatory cells.

### (iii) Herlev Pap smear dataset:

Herlev University Hospital and the Technical University of Denmark announced the Herlev Pap smear dataset (Jantzen et al. 2005) comprising of 917 Pap smear images, each of which encompasses one cervical cell segmented and classified by means of ground truth. The images in the dataset are captured at a magnification of  $0.201 \mu\text{m}/\text{pixel}$  with a resolution of  $156 \times 140$  pixels on average with the longest length of a side is 768 pixels and the shortest is 32 pixels. Seven classes of cell images are available in this dataset wherein first three classes namely superficial squamous, intermediate squamous, columnar are normal cells and remaining four classes namely are abnormal cells namely mild dysplasia, moderate dysplasia, severe dysplasia, and carcinoma in situ.

### (d) Microsoft common objects in context (MS COCO) dataset:

The MS COCO dataset (Lin et al. 2015) investigates the drawbacks of non-iconic views of object representation. Looking at Objects that are not the main emphasis of an image, is generally stated as a non-iconic view and this dataset was created with the help of Amazon Mechanical Turk for data annotations. MS COCO comprises of 2,500,000 labelled instances in 328,000 images and comprises of 91 common object categories, 82 of which have over 5,000 labelled instances.

### (v) Multi-organ nuclei segmentation (MoNuSeg) dataset:

Indian Institute of Technology Guwahati prepared a dataset named as **Multi-organ Nuclei Segmentation (MoNuSeg)** dataset that was published in the official satellite event of MICCAI 2018 contains WSI images of 7 organs (breast, kidney, colon, stomach, prostate, liver, and bladder) from various medical centres (i.e., various stains) of high-resolution WSI of H&E-stained slides from nine tissue types, digitised at  $40 \times$  magnification in eighteen different hospitals and obtained from National Cancer Institute's cancer Genome Atlas (TCGA) (Tomczak et al. 2015) with training set comprising of colour normalized (Vahadane et al. 2016) H&E images from all tissue types, excluding breast.

### (f) Data science bowl 2018 (DSB2018) dataset:

The DSB2018 (Caicedo et al. 2019) dataset is yet another dataset freely available at Broad Bio-image Benchmark

Collection (Data science bowl 2018) comprising of 670 images of segmented nuclei attained under diverse circumstances namely altering the cell type, magnification, and imaging modality (bright-field vs. fluorescence) that are further resized from various resolution to  $256 \times 256$  (aspect ratio maintained).

(vii) **Kasturba Medical College Liver (KMC Liver) dataset:**

The KMC Liver (Kasturba Medical College 2021) dataset containing 257 (70: sub-type 0; 80: sub-type 1, 83: sub-type 2, and 24: sub-type 3) original slides each measuring  $1920 \times 1440$  pixels, belonging to 4 sub-types of liver HCC tumour taken from various patients. It includes 80 H&E-stained histopathology images collected by pathologists at Kasturba Medical College (KMC), Manipal.

(viii) **PanNuke dataset:**

The PanNuke (Gamper et al. 2019) dataset comprise of H&E-stained image set containing 7,904 images of  $256 \times 256$  patches from 19 different tissue types wherein the nuclei are classified into 5 different categories of cell namely neoplastic, inflammatory, connective/soft tissue, dead, and epithelial cells. Gamper et al. (2020) outlines an evaluation process that separates the patches into three folds (later these 3 folds are used to create three different dataset splits) wherein one-fold is helpful for training and the remaining two folds for validation and testing sets, containing 2657, 2524, and 2723 images, respectively.

The following Fig. 2 shows a graphical representation of the most frequently used dataset over the last five years by the researchers, according to Table 1.

**4.3 Analysis based on optimizer**

Optimizer is a procedure of improving the neural network properties like weight and learning rates. It helps to minimizing the loss and enhancing performance. The following Fig. 3 shows a graphical representation of most frequently

used optimizers by the researchers according to Table 1. Adam is the widely used optimizer.

**4.4 Analysis based on loss function**

Loss Function is a process of examines how a CNN model predicts the intended results. The following Fig. 4 shows the most of the usable loss function that encountered while performing the literature survey as depicted in Table 1. BCE is the mostly used loss function.

**4.5 Analysis based on evaluation metric**

The evaluation metric or segmentation quality parameters are a measurement process of performance indicator for segmentation models. The following Fig. 5 shows some of the mostly used parameters that had been using the most of the work in the literature survey as depicted in Table 1.

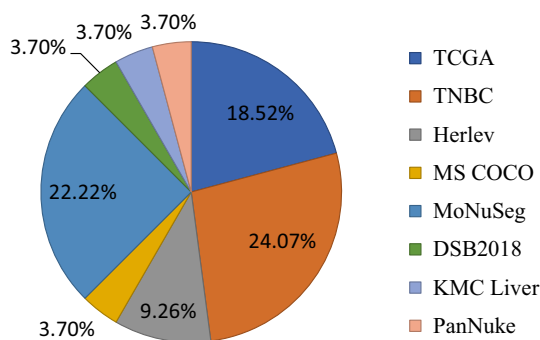


Fig. 2 Graphical representation of mostly utilised Datasets

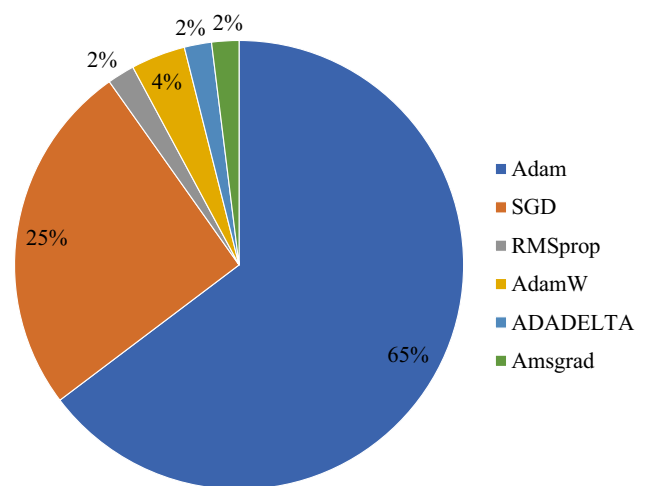


Fig. 3 Graphical representation of mostly utilised Optimizers

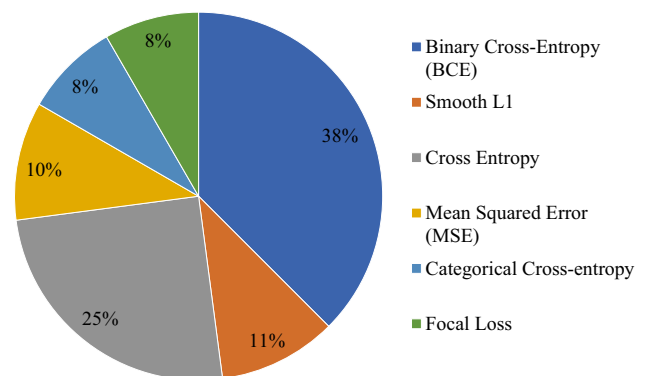


Fig. 4 Graphical representation of mostly utilised Loss Function

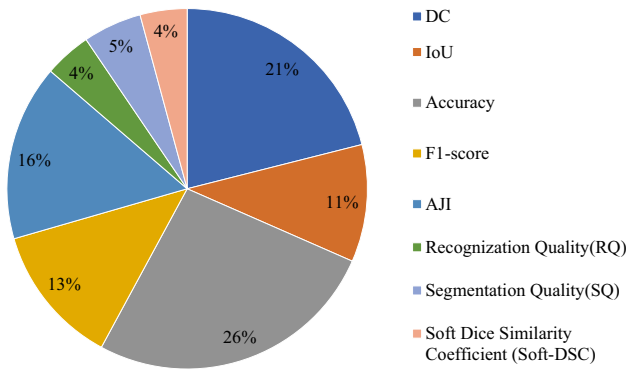


Fig. 5 Graphical representation of mostly utilised Evaluation Metric

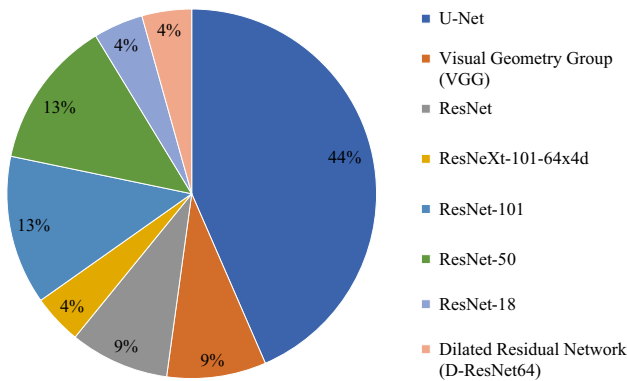


Fig. 6 Graphical representation of mostly utilised Backbone

### 4.6 Analysis based on backbone

Backbone means which feature extracting network is being used in the CNN model architecture. In the following Fig. 6 covers some of the backbone that has been used in Table 1

literature survey's models. U-Net is the most popular backbone for nucleus segmentation.

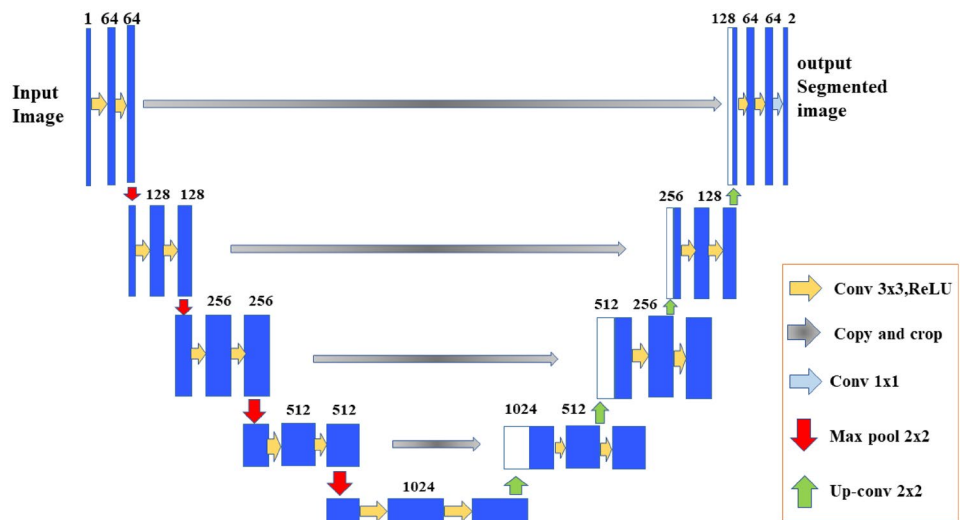
## 5 Experimental CNN models

This section provides an overview of some of the greatest CNN models that have been proposed up to this point, including U-Net, SCPP-Net, Sharp U-Net, and LiverNet. These are the models that we have utilized in our comparative analysis, and they are described in the following ways:

### 5.1 U-Net

FCNs and encoder-decoder models influenced numerous models originally designed for medical/biomedical picture segmentation. Qu et al. proposed the U-Net (Ronneberger et al. 2015) model, in which the network and training approach rely on data augmentation to effectively learn from a limited number of annotated images. The U-Net design, which is depicted in Fig. 7, comprises two parts: a contracting path for context capture and a symmetrically extending path for accurate localization. An FCN-like design extracts features with  $3 \times 3$  convolutions in the down-sampling or contracting section. Up-convolution, popularly known as deconvolution, is used for feature map reduction and up-sampling for increasing the dimensions to prevent losing pattern information. Navigation of feature maps takes place from the network's down-sampling section towards the up-sampling section. Finally, feature map analysis takes place with the help of a  $1 \times 1$  convolution, further creating a segmentation map that classifies each pixel in the input picture. For different types of pictures, several U-Net extensions have been developed. A multi-channel feature map is represented by each blue box in Fig. 7 with channel numbers on top, and the white boxes represent the copies of the

Fig. 7 Architecture of U-Net Model



feature map. The sizes X and Y are indicated in the lower left border of the box, whereas the arrows represent the various operations being carried out.

## 5.2 Separable convolutional pyramid pooling network (SCPP-Net)

This SCPP-Net model by Chanchal et al. (2021b) was built upon the idea of mining supplementary information at an advanced level, as depicted in Fig. 8. The receptive field of the SCCP layer is expanded by keeping the kernel size constant while regulating four distinct dilation rates. The generated feature maps have an extra parameter called “dilation rate” that could be changed to see bigger areas. The separation of clumped and overlapping nuclei is a critical issue in histopathology image nuclei segmentation. However, by expanding the receptive field at a higher level, this

CNN-based design helps to overcome the problem of proximity and overlapping nuclei.

The convolution and max-pooling operations are conducted on the input image during the down-sampling process, giving extreme importance to capturing the context of the image, which leads to the growth in the size of the image on the one hand but, on the other hand, depth drops along the growing route. For the same reason, progressively adding up-sampling to the decoder route enables accurate localization. Figure 8 depicts the proposed SCPP-Net’s comprehensive design, whereas Fig. 9 depicts the SCPP-Net’s inclusive and precise SCPP block concept.

## 5.3 Sharp U-Net

In encoder-decoder networks, predominantly in U-Net (Ronneberger et al. 2015), for the purpose of convalescing fine-grained features, skip connections play a vital role for

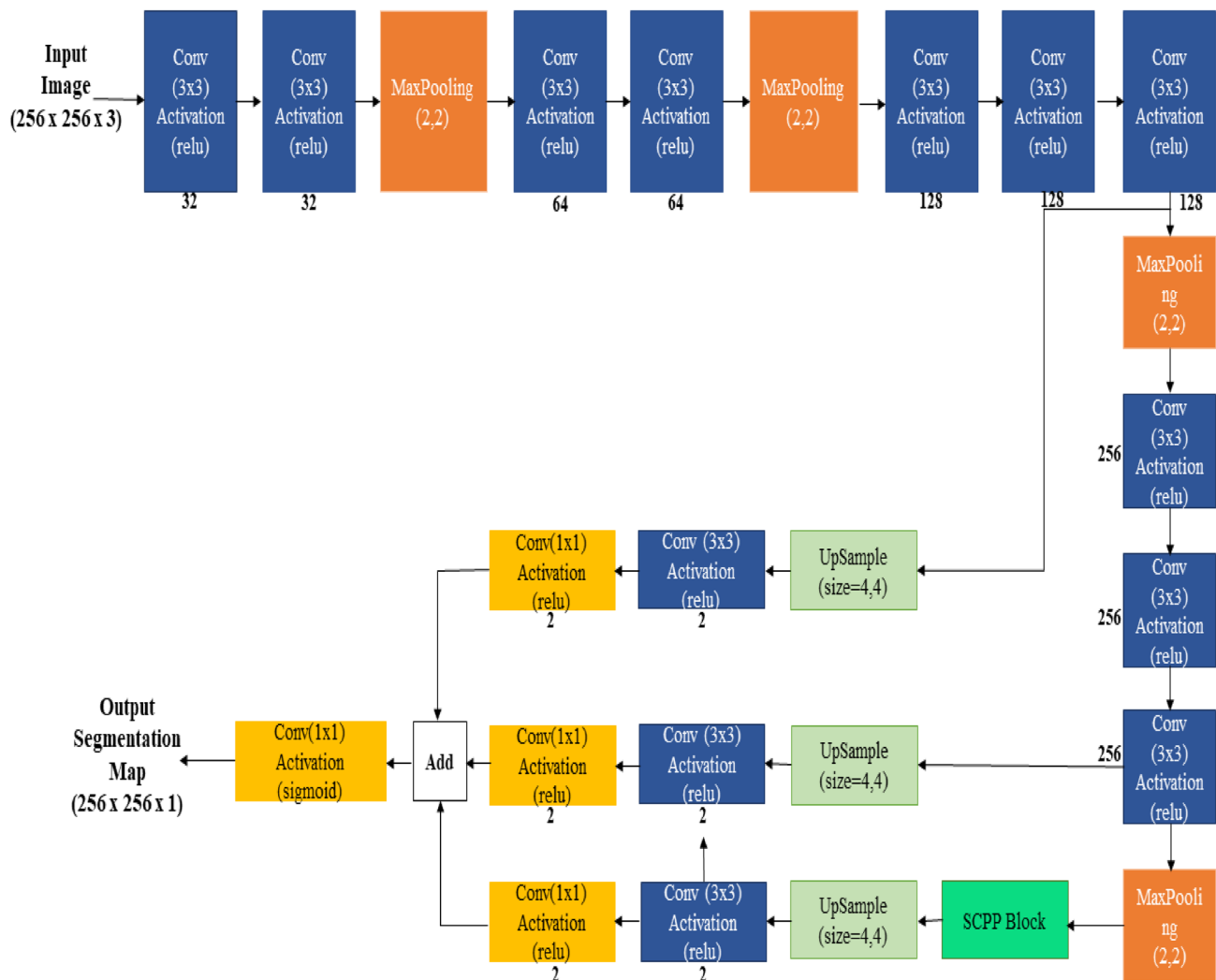
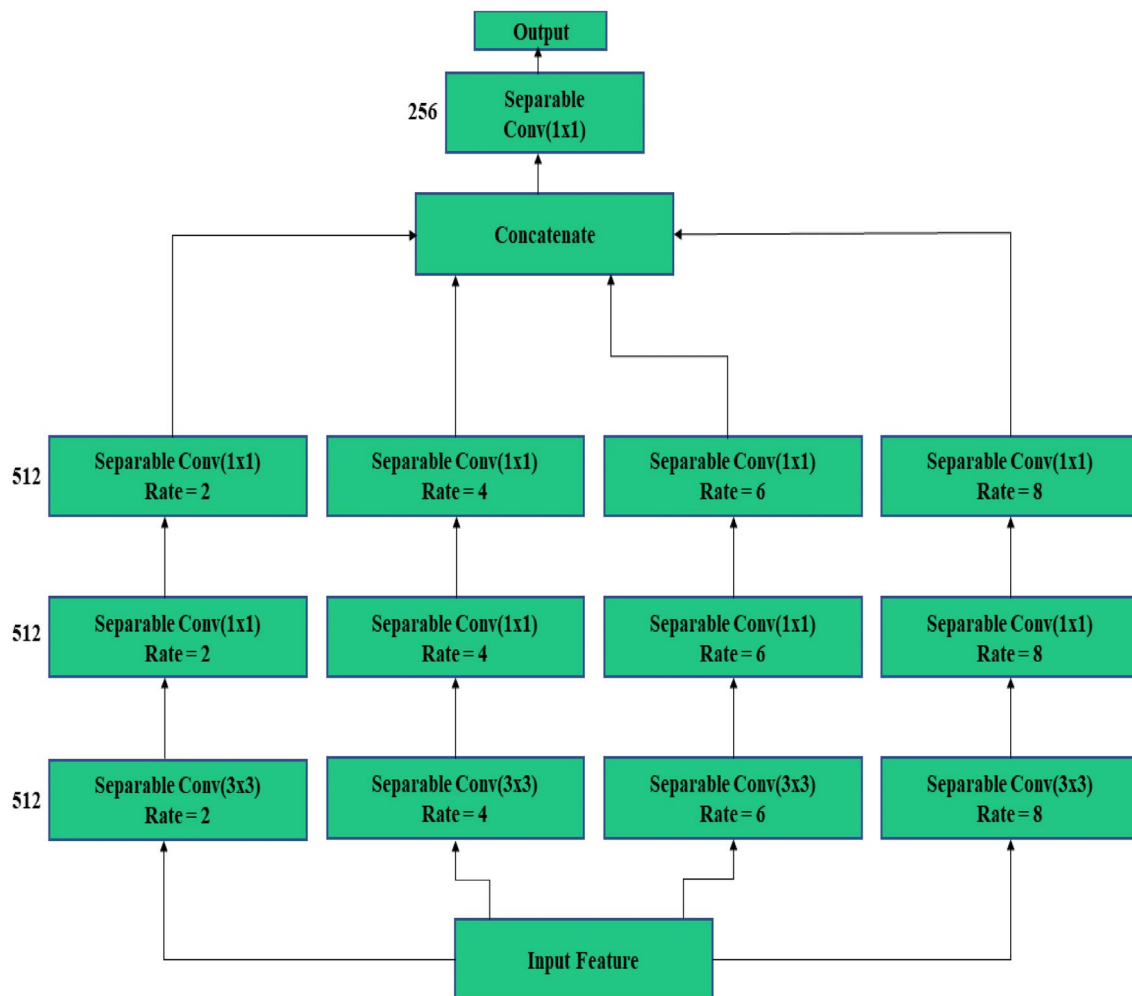


Fig. 8 Architecture of Separable Convolution Pyramid Pooling Network (SCPP-Net)



**Fig. 9** Separable Convolution Pyramid Pooling (SCPP) block

prediction. Moreover, skip connections have a tendency to semantically associate low- and high-level convolution features of diverse nature, thereby generating totally obscure feature maps. In order to overcome such a flaw, Zunair and Hamza et al. suggested the Sharp U-Net (Zunair and Hamza 2021) architecture, as revealed in Fig. 10 that is applicable to both binary and multi-class segmentation.

The encoder section is divided into five blocks, each of which includes two  $3 \times 3$  convolutional layers with ReLU activations, followed by a  $2 \times 2$  layer known as a “max-pooling layer.” For the convolutional layers, 32, 64, 128, 256, and 512 filters are applied, and the same are used along the input to construct a feature map that basically recapitulates the occurrence and existence of the features that have been mined from the said input. A new connection mechanism, termed a “sharp block,” as depicted in Fig. 11, is formed to contain the up-sampled features with the intention of fusing the encoder and decoder’s low- and high-level features, avoiding the semantic gap issues. The encoder features are

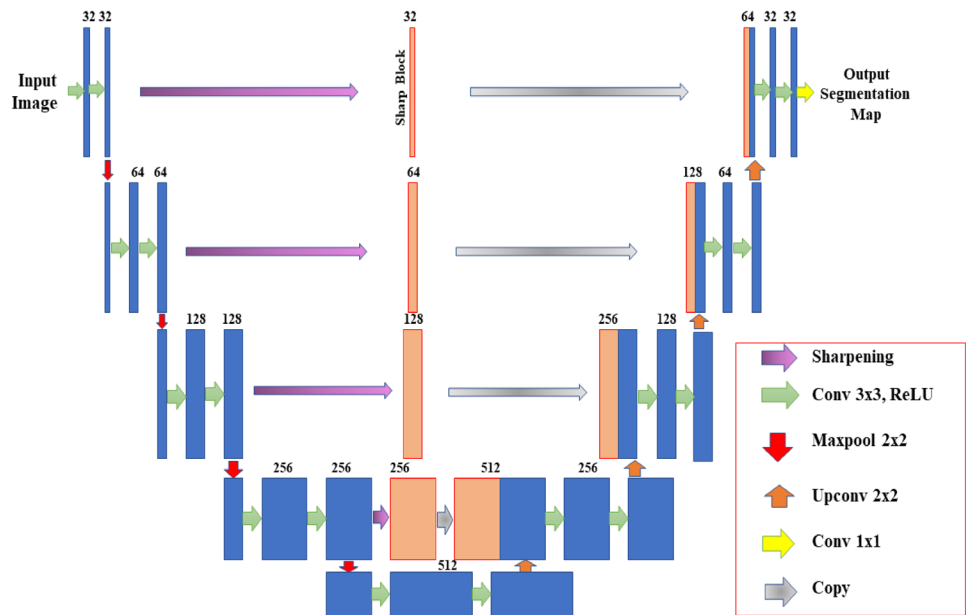
rather exposed to a spatial convolution operation that is fundamentally accomplished autonomously on each channel of the encoder features by means of a sharpening spatial kernel beforehand and then making use of meek skip connections between encoder and decoder.

#### (a) Sharpening Spatial Kernel

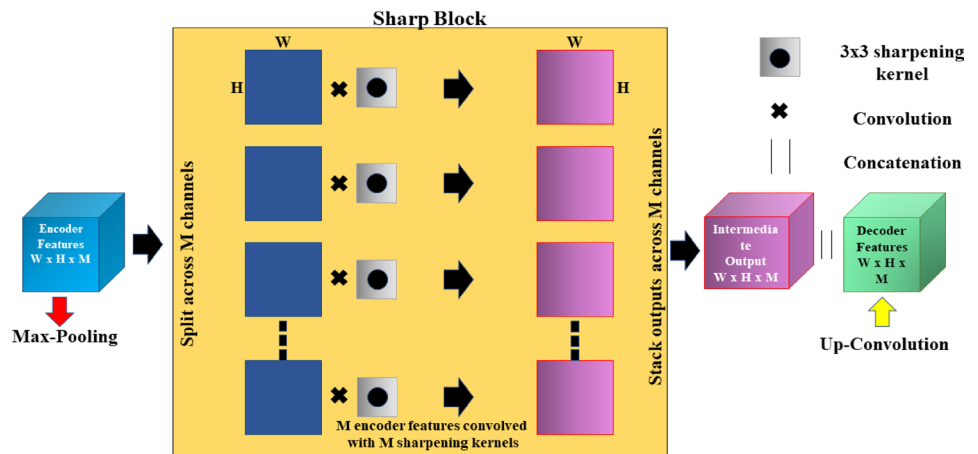
Spatial filtering, on the other hand, is a low-level neighbourhood-based image processing method that basically tends to enhance the image (sharpen the image) by performing certain operations on the neighbourhood of individual pixels of the input image. Image convolution with kernels is used to perform high-pass filtering or image sharpening. Convolution kernels, normally referred to as filters, are a second-order derivative operator that might respond to intensity evolutions in any direction. A typical Laplacian high-pass filtering kernel is specified as a matrix,  $K$ , that includes a negative value off-center and a single positive value in the



**Fig. 10** Architecture of Sharp U-Net



**Fig. 11** Illustration of Sharp Block



centre for image sharpening that takes into account all eight neighbours of the input image's reference pixel.

$$K = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Kernel adjusts the brightness of the centre pixel in relation to the adjacent pixels while convolving an image with the Laplacian filter. Additionally, the input imagery is added to its convolution with the kernel to produce a refined image. Considering an input image,  $I$ , and the resultant sharpened image  $S$ ,  $S$  is generated as  $S = I + K * I$ ; wherein  $*$  signifies convolution, a kernel weighted neighbourhood-based operator that processes an image by adding each pixel's value to its nearest neighbours.

(b) **Sharp block**

This block does a depth-wise convolution on a single feature map by using a sharpening spatial kernel given by the Laplacian filter kernel  $K$ . This kernel is usually of size  $W \times H \times M$ , where  $W$ ,  $H$ , and  $M$  are the width, height, and number of the encoder's feature maps, respectively.

In convolutions,  $M$  filters are applied that discretely act on each of the input channels rather than a single filter of a specific size (i.e., 333). Individual input channels are convolved with the kernel  $K$  individually with a stride of 1, thereby producing a feature map of  $W \times H \times 1$  dimension. To retain the dimension of the output to be the same as that of the input and to match the size of the decoder features all throughout the connection, padding is performed during the feature fusion of the encoder and decoder sub-networks.

The depth-wise convolution layer's ultimate output of size  $W \times H \times M$  is attained at this point by piling these maps together. This planned feature connection is referred to as a “sharp block.” Fig. 11 displays a visual representation of the sharp block's operation flow.

### 5.4 LiverNet

The convolution procedure resides in the heart of every CNN. 2D discrete linear convolution is articulated as (1) with  $f$  and  $h$  as two-dimensional signals. Aatresh et al. (2021a) suggested the LiverNet model for liver hepatocellular carcinoma histopathology images.

$$g(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f[k, l]h[m - k, n - l] \tag{1}$$

Using the above definition, they add the bias to Eq. (1) to obtain the computation formula per node in a given layer. In addition, Max-pooling operations are a critical operation in most CNN systems nowadays (Krizhevsky et al. 2017). Consider a sliding window over the input feature map to the max-pool layer to better understand this procedure. By sliding the window with a stride  $S$ , this procedure provides the greatest value of the pixels inside the window, which is repeated throughout the entire image. By lowering the number of parameters, max-pool layers help minimise the network's computational complexity and provide an abstract representation of the input data.

Aatresh et al. (2021a) employ a base architecture similar to Toğaçar et al. (2020), and they extract features from the input image using two convolution layers before the initial max-pool operation. To extract relevant information more effectively, they used CBAM

blocks (Woo et al. 2018) and residual blocks deeper in the architecture. After each max-pool operation, they employ intermediate features in the encoder pipeline to feed into ASPP blocks before up-sampling. To merge the pixel data of layers of varied depths, the hyper-column approach employed in Toğaçar et al. (2020) was applied. The hyper-column technique, along with ASPP blocks, ensures multi-scale feature extraction and information retrieval for further processing in this architecture. They have applied these ideas to the problem of multi-class cancer classification in liver tissue, and a detailed depiction of the proposed model can be found in Fig. 12. The sub-modules of the LiverNet architecture have been described in detail in the following subsections.

#### (a) CBAM block and residual block

Convolutional Block Attention Module (CBAM), introduced by Woo et al. (2018), comprised of a CBAM block that is proficiently implanted into any CNN architecture without instigating unnecessary computation or memory performance drawbacks. Channel-wise and spatial attention modules were anticipated in succession to produce attention maps that were increased by the input feature map. In CBAM, the channel-wise attention block focuses on **what** the network needs to focus on, whereas the spatial attention block concentrates on **where** the network needs to place emphasis.

The CBAM block's behaviour at an intermediate step, taking into consideration a feature map  $A \in \mathbb{R}^{H \times W \times C}$  input in the encoder pipeline, can be mathematically projected as in (2).

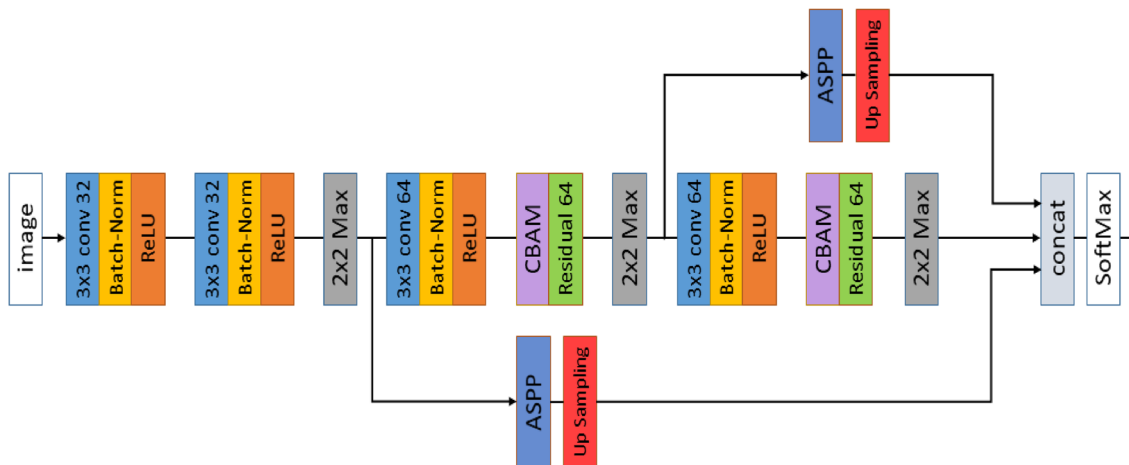


Fig. 12 Architecture of LiverNet

$$A_c = f_c(A) \cdot A(2) \quad \text{and} \quad A_s = f_s(A_c) \cdot A_c \quad (2)$$

where “.” representing element-wise multiplication;  $f_c: \mathbb{R}^{H \times W \times C} \rightarrow \mathbb{R}^{1 \times 1 \times C}$  and  $f_s: \mathbb{R}^{H \times W \times C} \rightarrow \mathbb{R}^{H \times W \times 1}$  symbolizing the functions of channel-wise and spatial attention blocks, correspondingly. Following the element-wise multiplication between the channel-wise attention map  $f_c$  and the input feature map  $A$ ,  $A_c$  is the intermediate output. The product of element-wise multiplication between the spatial attention map and  $A_c$ , as well as the final output of the CBAM attention block. The channel-wise attention block is composed of concurrent average and max-pooling procedures that share a fully connected network as described in Eq. (3) before it is added.

$$f_c = \sigma(FC(AvgPool(A))) + FC(MaxPool(A)) \quad (3)$$

wherein  $\sigma$  is the popular sigmoid function with FC being the shared fully connected layers. Before feeding the result to a convolution layer, the spatial attention block concatenates the results of max-pool and average-pool operations. Whenever the input  $A$  is provided, the action is defined by Eq. (4).

$$f_c = \sigma(w \otimes ([AvgPool(A); MaxPool(A)])) \quad (4)$$

wherein  $\otimes$  represents the two-dimensional convolution operation with a kernel  $w$ . (He et al. 2016) has proposed the residual block that is used in the LiverNet architecture, which is comparable to the residual block used in Toğaçar et al. (2020). The main difference is that the filters in the residual block’s initial convolution layer are lowered by a factor of 4 when compared to the filter used in the residual block presented in Toğaçar et al. (2020). This not only reduced the number of parameters needed in the model but it also increased the quality of the features derived from the input.

(b) **ASPP block**

An Atrous Spatial Pyramid Pooling (ASPP) block may successfully extract multi-scale features from a feature map, as demonstrated in Chen et al. (2018). They use a comparable ASPP block in the LiverNet architecture because of its effectiveness. To increase the size of the receptive field without increasing the number of parameters involved, atrous convolution or dilated convolution can be utilised. Consider a two-dimensional signal  $X$  that has been convolved with a two-dimensional filter  $w$  via Atrous convolution. A convoluted product is represented by the following equation Eq. (5).

$$Y[i, j] = \sum_m \sum_n X[i + r.m, j + r.n] w[m][n] \quad (5)$$

where  $r$  corresponds to the dilation rate or the rate at which the input signal  $X$  is sampled. Atrous convolution has the

effect of increasing the receptive field size of the kernel by adding  $r-1$  zeros in between the kernel elements. As a result, if  $r=2$ , a  $3 \times 3$  kernel will have some receptive field size equivalent to a  $5 \times 5$  kernel but with just 9 parameters.

Figure 13 illustrates the ASPP block that is employed in the LiverNet architecture. A feature map is received as an input before concatenation thereafter certain operations are conducted in parallel namely  $1 \times 1$  Convolution;  $3 \times 3$  Convolution with dilation rate = 2;  $3 \times 3$  Convolution with dilation rate = 3;  $3 \times 3$  Convolution with dilation rate = 6;  $3 \times 3$  Convolution with dilation rate = 8 and global average pooling.

To keep the same filter size as the input, the entire convolution and pooling outputs are concatenated and passed through a 11 convolution layer. Further, the convolution output is passed through a batch normalization and ReLU activation layer before being delivered to the bilinear up-sampling layer. The output of the max-pool layers delivers feature-rich information at many sizes and extents; therefore, the ASPP block is placed after each max-pooling operation in the encoder pipeline in the LiverNet architecture.

For the entire models used in our work, we used Binary Cross Entropy (BCE) as the loss function, as well as the Intersection over Union (IoU) and Dice Coefficient (DC) parameters for the quantitative analysis of the nucleus segmentation results. We define these loss functions and parameters as follows:

(i) **Loss function**

Reducing the loss is the goal of an error-driven learning algorithm, which is accomplished through the use of a good loss function. We anticipated a number and were eager to find how much we were off, the squared error loss appears appropriate for the regression problem. We recognize it was a distribution for classification, so we could use something that captures the difference between the true and projected distributions. In our study, we use a loss function named Binary Cross-Entropy (BCE) (Ahamed et al. 2020).

$$\text{Cross Entropy Loss} = - \sum_{k=1}^C y_k \log(f(s)_k) \quad (6)$$

where  $y_k$  and  $s_k$  represent the ground truth and projected scores for each class  $k$  in  $C$  respectively. For loss computation, ReLU activation in the intermediate layer and sigmoid activation before are used.. Two distinct symbols,  $C$  and  $C'$ , represent two classes as used in different equations wherein for  $C$  classes, Eq. (6) represents cross-entropy loss; for  $C'$  classes, Eq. (7) represents BCE loss. As projected in Eq. (8) BCE loss is highlighted with respect to the activation unit is either denoted as  $f(s_k)$  or  $\hat{y}$ .

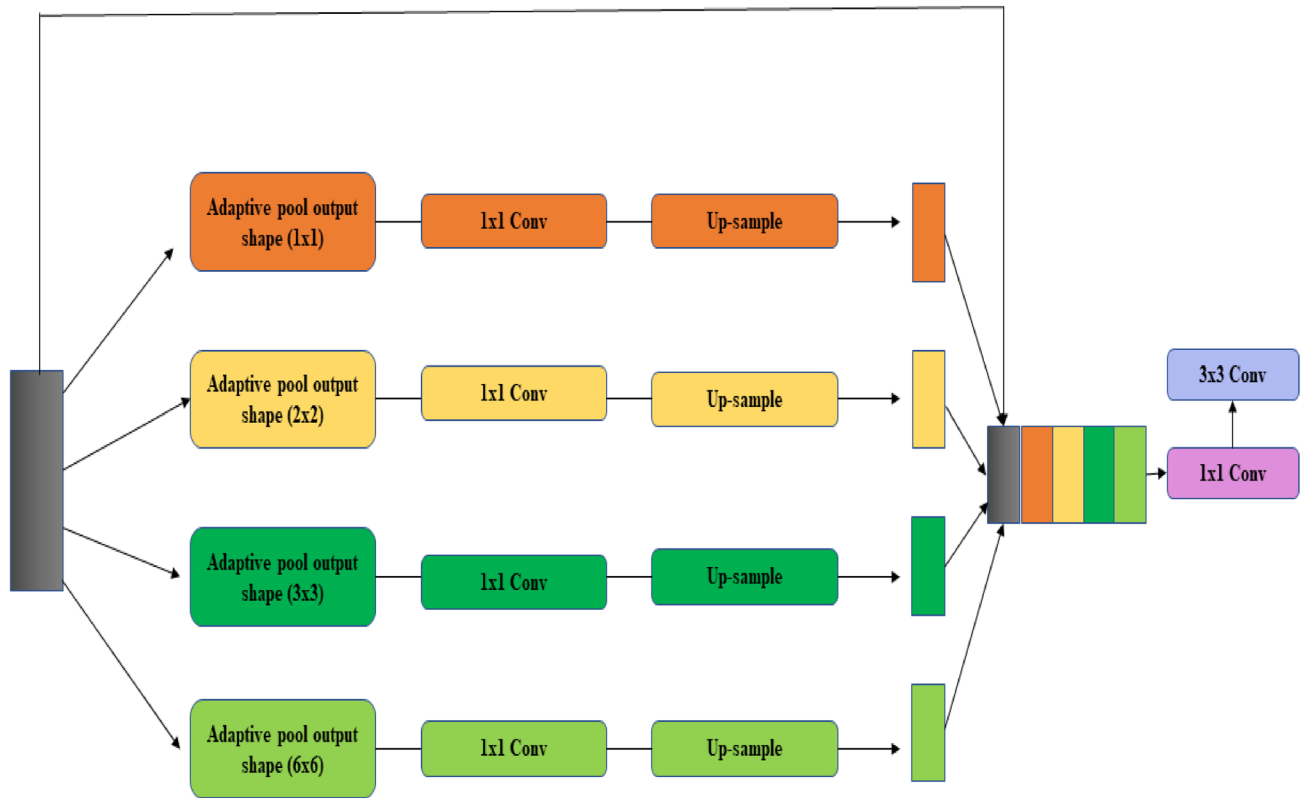


Fig. 13 ASPP Block in LiverNet Architecture

$$\text{BCE Loss} = - \sum_{k=1}^{c'} y_k \log(f(s)_k) \tag{7}$$

$$\text{BCE Loss} = -[y \ln(\hat{y}) + (1 - y) \ln(1 - \hat{y})] \tag{8}$$

(b) Segmentation quality parameters

In purpose of our study, two segmentation quality parameters have been used, such as Intersection over Union (IoU) (Kanadath et al. 2021) and Dice Coefficient (DC) (Gudhe et al. 2021).

(A) **Intersection over Union (IoU):** In the fields of semantic segmentation, IoU, popularly known as the Jaccard Index, is yet another frequently used metric that is basically the area overlapped between predicted segmentation and the ground truth, as indicated by the area of union between the predicted segmentation and the ground truth indicated in Eq. (9), wherein A is the ground truth mask image and B is the predicted segmentation result obtained from the model.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{|A \cap B|}{|A \cup B|} \tag{9}$$

(B) **Dice coefficient (DC):** This segmentation quality parameter measures the similarity between the predicted mask and the corresponding ground truth mask, which is generally defined as 2 multiplied by the area of overlap divided by the total number of pixels in both images, as depicted in Eq. (10), wherein A is the ground truth mask image and B is the predicted segmentation result obtained from the model.

$$\text{DC} = \frac{2 \times \text{Area of Overlap}}{\text{Total Number of pixels}} = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{10}$$

## 6 Experimental result and discussion

This section represents the experimental results of the four well-known deep learning CNN models, namely U-Net, Separable Convolutional Pyramid Pooling Network (SCPP-Net), Sharp U-Net, and LiverNet, over a merged dataset. The specifics of the dataset we used are briefly described below.

## 6.1 Experimental dataset

For our purpose, we merged three publicly available datasets, such as JPI 2016 dataset (Janowczyk and Madabhushi 2016), IEEE TMI 2019 dataset (Naylor et al. 2018) and PSB 2015 dataset (Irshad et al. 2015), respectively. These three datasets are described in details as follows:

- (A) **JPI 2016 Dataset:** Janowczyk and Madabhushi (2016) announced this dataset which comprises of 143 H&E images of 137 patients and ER + BC a images scanned at 40x. Each image is 2000 by 2000 pixels in size, with around 12,000 nuclei painstakingly segmented throughout the photos. The file is in the following formats: 12750\_500\_f00003\_original.tif for original H&E photos and 12750\_500\_f00003\_mask.png for a mask of the same size, with white pixels representing nuclei. Each image is prefaced by a code i.e., 12,750 to the left of the first underscore (\_), which defined with a unique patient number. A few patients (137 patients vs. 143 images) have several images associated with them.
- (B) **IEEE TMI 2019 Dataset:** Naylor et al. (2018) offered this IEEE TMI 2019 dataset generated by the Curie Institute, which comprises of annotated H&E-stained histology images at 40× magnification wherein total of 122 histopathology slides are annotated. There are 56 annotated pCR, 10 are RCB-I, 49 are RCB-II and 7 are RCB-III.
- (C) **PSB 2015 Dataset:** Irshad et al. (2015) presence this PSB 2015 dataset, images in this dataset came from the TCGA data portal's WSIs of Kidney Renal Clear Cell Carcinoma (KIRC). The TCGA project is jointly supported by National Cancer Institute and the National Human Genome Research Institute and TCGA has undertaken detailed molecular profiling on tens of thousands of tumours, covering the 25 most frequent cancer types. 10 KIRC Whole Slide Images (WSI) from the TCGA data portal (<https://tcgadata.nci.nih.gov/tcga/>) is selected. Nucleus-rich ROIs and extracted 256×256-pixel size images for each ROI at 40× magnification is identified further from these WSIs.

Therefore, there are a total of 653 images contained inside the combined dataset that was employed. We used random selection to choose 457 images from these three datasets to use for training, 98 images to use for validation, and 98 images to use for testing.

## 6.2 Training and implementation

To speed up the development procedure and experiments on a machine with Ryzen 5 3550, 16 GB RAM, and a Nvidia GTX 1650, the training and implementation were done in a

Jupyter notebook with the latest version of Keras and TensorFlow Python-3 framework. The four deep learning models considered in this study were trained using Sigmoid or SoftMax as the activation function and an adaptive learning rate optimization algorithm known as the Adam optimizer to speed up the training. The loss function employed for the four models is binary cross entropy (BCE) (Ahamed et al. 2020), as highlighted in (9). Further, batch sizes of 8, 4, 10, and 2 for U-Net, SCPP-Net, Sharp U-Net, and LiverNet, respectively, are used to train all 256×256 histopathology images.

## 6.3 Discussion on segmentation results

In this study, a comparative examination of four pre-trained CNN architectures—U-Net (Ronneberger et al. 2015), SCPP-Net (Chanchal et al. 2021b), Sharp U-Net (Zunair and Hamza 2021), and LiverNet (Aatresh et al. 2021a)—is conducted. On a combined dataset, all four models are trained using almost 457 images for training, 98 for validation, and 98 for testing. During training, the network is fed the histopathological images from the training set and the ground truth masks. Two assessment metrics are proposed in this study, namely intersection over union (IoU) (Kanadath et al. 2021) and dice coefficient (DC) (Gudhe et al. 2021), which are shown in (9) and (10), respectively. All the models are then used to predict the masks of the test images. The input size for all models is 256×256 pixels. The U-Net and SCPP-Net models that were used had, respectively, 7,725,249 and 2,985,659 trainable parameters. The Sharp U-Net and LiverNet models, on the other hand, have 4,320 and 12,288 non-trainable parameters, respectively, and 7,760,097 and 989,117 trainable parameters. The training times for the U-Net, SCPP-Net, Sharp U-Net, and LiverNet models, which are trained over 1500, 500, 550, and 700 epochs, respectively, are 590 ms, 330 ms, 595 ms, and 490 ms per step. Because U-Net is less sophisticated than the other three models, it takes slightly less time. On the other hand, Sharp U-Net provides better image segmentation and accuracy. The performance of four deep learning models for nuclear segmentation (U-Net, SCPP-Net, Sharp U-Net, and LiverNet) is fairly compared in Table 2.

The main model complexity of U-Net (Ronneberger et al. 2015) architecture is that the resulting segmentation map will be negatively impacted by the feature mismatch in between encoder and decoder paths, which will cause the fusing of semantically incompatible data and hazy feature maps during the learning process. The segmentation of overlapping nuclei is the main complexity of the SCPP-Net (Chanchal et al. 2021b) model. Sharp U-Net (Zunair and Hamza 2021) predicts segmented outcomes that are slightly under-segmented and defective, but it generates far less noise and segmented outputs that are broken. The key



**Table 2** Performance comparison of four utilised architectures on merged dataset

Model	IoU	Dice	Accuracy (%)
U-Net	0.4934	0.6599	<b>83.13</b>
SCPP-Net	0.4711	0.6340	81.67
Sharp U-Net	<b>0.5276</b>	<b>0.6899</b>	82.04
LiverNet	0.3810	0.5299	82.28

Best results are highlighted in bold

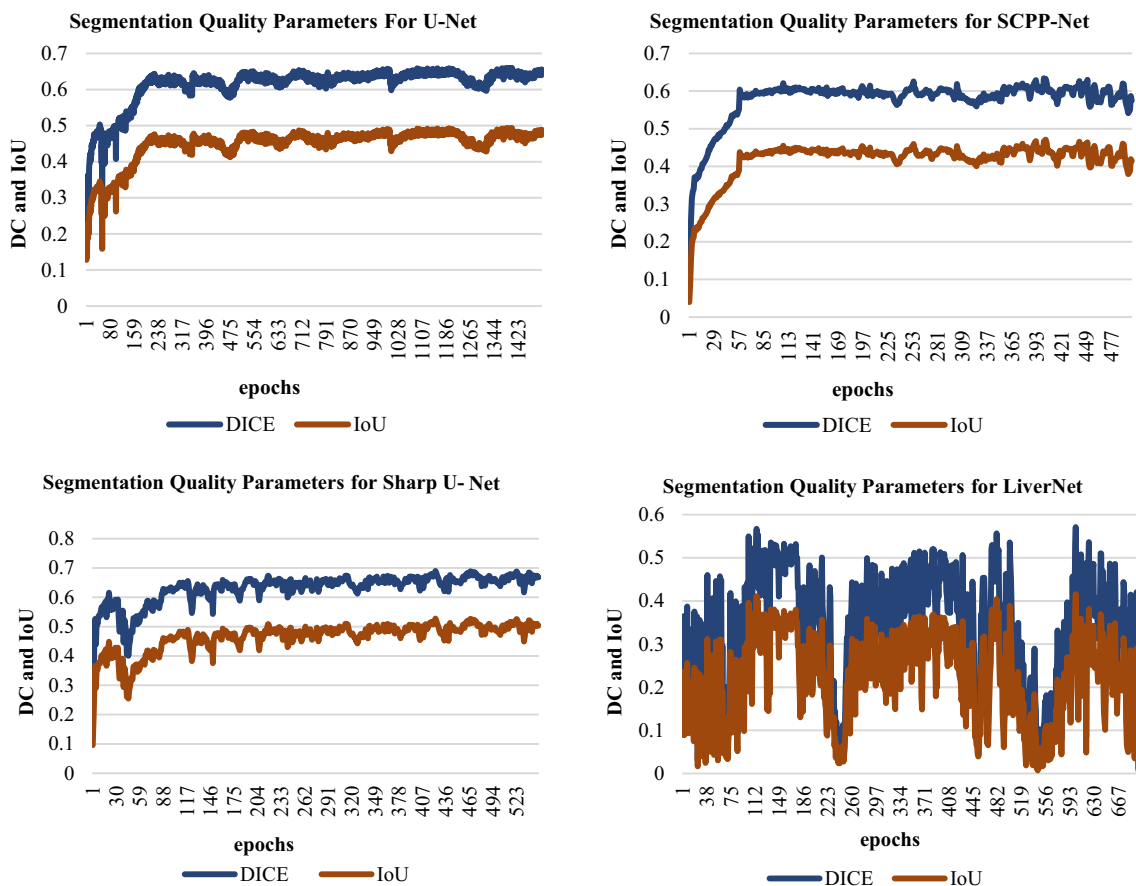
difficulty of the LiverNet (Aatresh et al. 2021a) model is that it was difficult to segment the tiniest and most densely packed nuclei.

Figure 14 depicts a graphical representation of the segmentation experiment based on (IoU Score, Dice, and Accuracy%), which clearly demonstrates that Sharp U-Net produces the best segmentation results for two quality parameters, namely IoU and Dice, producing smoother predictions than the other three segmentation models used.

In terms of the Dice Coefficient (DC) and Intersection over Union (IoU) score, the segmentation results of four nuclei segmentation models considered the merged

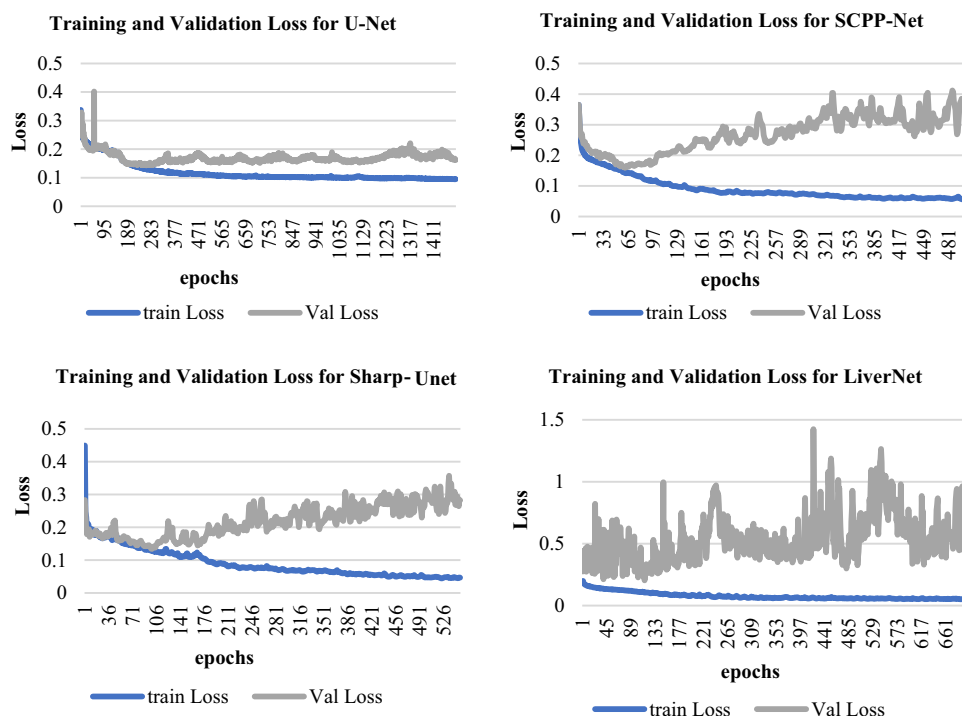
dataset. As shown, LiverNet obtains a DC of 0.5299 and an IoU of 0.3801, which are lower than the other three models. U-Net and SCPP-Net achieve improvements on the DC and IoU scores. U-Net and SCPP-Net obtain DC = 0.6599 and IoU = 0.4934 and 0.4711, respectively, as depicted in Table 2. Sharp U-Net obtains the best results on the DC and IoU at 0.6899 and 0.5276, respectively. This analysis further reveals that Sharp U-Net could be used to get suitable nuclear segmentation results. The four segmentation models used (U-Net, SCPP-Net, Sharp U-Net, and LiverNet) produce accuracy of 83.13%, 81.67%, 82.04%, and 82.28%, respectively. Figure 15 depicts a graphical representation of the training and validation loss for four CNN models.

Figure 16 contains examples of some original images as well as images predicted by various models. These examples highlight the various outcomes of our segmentation results based on our combined dataset. Based on the information shown in this figure, the Sharp U-Net produces a better segmented image than the other three models that were tested.



**Fig. 14** Graphical representation of IoU and DC of our four utilised segmentation models

**Fig. 15** Graphical representation of training and validation loss of four utilised CNN models



## 7 Conclusion and future directions

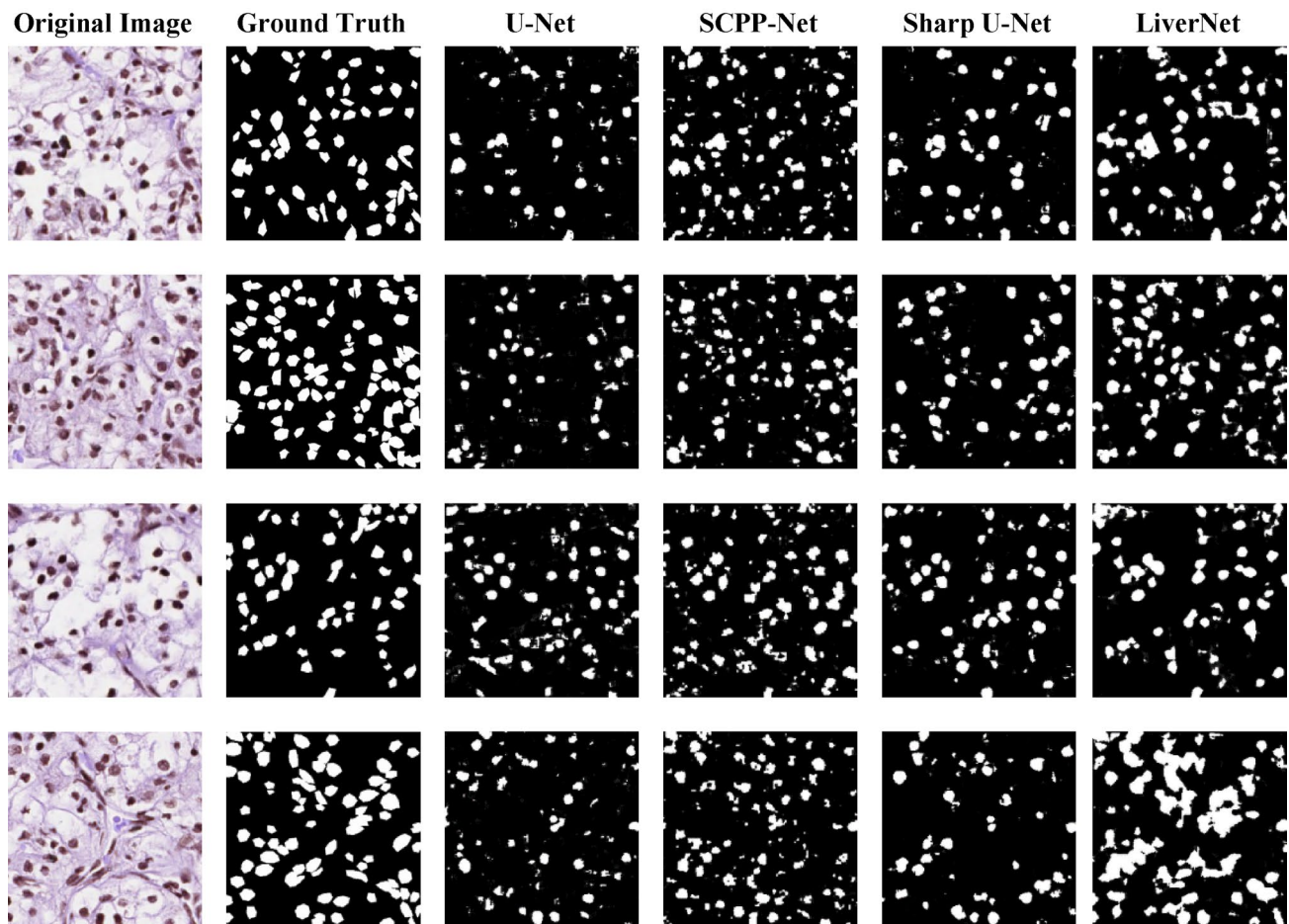
Recent advancements in the field of computer vision and machine learning strengthen an assemblage of algorithms with remarkable and noteworthy ability to interpret the content of imagery. Several such deep learning algorithms are being imposed and employed on biological images, thereby massively transmuting the analysis and interpretation of imaging data and generating satisfactory outcomes for segmentation and even classification of images across numerous domains. Even though learning parameters in deep architectures necessitate a large volume of labeled training data, transfer learning is promising in such scenarios because it focuses on reusing the learned features and applying them appropriately based on the situation's requirements and demands. This study has three major contributions as a survey paper, which is stated below:

- An overview table of deep learning models used for nucleus segmentation from 2017 to 2021, with different optimizers used across a range of datasets and for different types of images, will show how different deep learning models are used for nucleus segmentation.
- A study that makes comparisons between four different deep learning models that were developed very recently for segmenting nuclei.
- Training the deep learning models mentioned in (ii) was performed by the merged version of three datasets, namely JPI 2016, IEEE TMI 2019, and PSB 2015, each containing thousands of images, and the training

results are updated in Table 2 and grouped according to the accuracy results. The experimental results are very encouraging; highlighting that Sharp U-Net delivers high accuracy results in all the cases with minimal loss. The highest accuracy obtained by Sharp U-Net is a DC of 0.6899 and an IoU of 0.5276. The DC and IoU values for U-Net, SCPP-Net, and LiverNet were 0.6599 and 0.4934, 0.6340 and 0.4711, and 0.5299 and 0.3801, respectively.

Therefore, it can be easy to infer that deep learning-based nucleus segmentation for histopathology images is a fresh and exciting research topic to work on and concentrate on. The major challenges one might encounter in the future would be to develop:

- Innovative and hybrid CNN architectures should be enabled for a wide range of medical image segmentation techniques,
- loss function should be designed for more specific medical image segmentation,
- The researchers should place a strong emphasis on transfer learning as well as the interpretability of the CNN model.
- Nature Inspired Optimization Algorithms (NIOA) based optimized CNN models should be explored.
- Explore different techniques and Architectures will be explored to further improve the speed and decrease the model size. in addition, larger diversified silent object



**Fig. 16** Row-wise visual segmentation comparison of four utilised models on merged dataset

datasets are needed to train more accurate and robust models

- f. Future study will focus on developing deep architecture that requires fewer calculations and can work on embedded devices while producing better test results.
- g. The information recession problem needs to be effectively mitigated that occurs in traditional u-shape architecture.
- h. Nature inspired optimization algorithms (Rai et al. 2022) like Aquila Optimizer (Abualigah et al. 2021a), Reptile Search algorithm (Abualigah et al. 2022), and Arithmetic optimization algorithm (Abualigah et al. 2021b) can be utilized to build optimized CNN models in the field of medical image segmentation.

**Funding** There is no funding related with this research.

**Data availability** The authors do not have the permission to share the data.

## Declarations

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest. The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Aatresh AA, Alabhya K, Lal S, Kini J, Saxena PP (2021a) LiverNet: efficient and robust deep learning model for automatic diagnosis of sub-types of liver hepatocellular carcinoma cancer from H&E stained liver histopathology images. *Int J Comput Assist Radiol Surg.* <https://doi.org/10.1007/s11548-021-02410-4>
- Aatresh AA, Yatgiri RP, Chanchal AK, Kumar A, Ravi A, Das D et al (2021b) Efficient deep learning architecture with dimension-wise pyramid pooling for nuclei segmentation of histopathology images. *Comput Med Imaging Graph* 93:101975
- Abdolhoseini M, Kluge MG, Walker FR, Johnson SJ (2019) Segmentation of heavily clustered nuclei from histopathological images. *Sci Rep* 9(1):1–13

- Abualigah L, Yousri D, Abd Elaziz M, Ewees AA, Al-Qaness MA, Gandomi AH (2021a) Aquila optimizer: a novel meta-heuristic optimization algorithm. *Comput Ind Eng* 157:107250
- Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021b) The arithmetic optimization algorithm. *Comput Methods Appl Mech Eng* 376:113609
- Abualigah L, Abd Elaziz M, Sumari P, Geem ZW, Gandomi AH (2022) Reptile Search Algorithm (RSA): a nature-inspired meta-heuristic optimizer. *Expert Syst Appl* 191:116158
- Ahamed MA, Hossain MA, Al Mamun M (2020) Semantic segmentation of self-supervised dataset and medical images using combination of u-net and neural ordinary differential equations. In: 2020 IEEE Region 10 symposium (TENSYP), pp 238–24
- Ahmed L, Iqbal MM, Aldabbas H, Khalid S, Saleem Y, Saeed S (2020) Images data practices for semantic segmentation of breast cancer using deep neural network. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-020-01680-1>
- Akram SU et al (2018) Leveraging unlabeled whole-slide-images for mitosis detection. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 11039 LNCS, pp 69–77
- Ali MA, Misko O, Salumaa SO, Papkov M, Palo K, Fishman D, Parts L (2021) Evaluating very deep convolutional neural networks for nucleus segmentation from brightfield cell microscopy images. *SLAS DISCOV: Adv Sci Drug Discov* 26:1125–1137
- Allehaibi KHS, Nugroho LE, Lazuardi L, Prabuwo AS, Mantoro T (2019) Segmentation and classification of cervical cells using deep learning. *IEEE Access* 7:116925–116941
- Alom ZMd, Aspiras TH, Taha TM, Asari VK, Bowen TJ, Billiter D, Arkell S (2019) Advanced deep convolutional neural network approaches for digital pathology image analysis: a comprehensive evaluation with different use cases. *CoRR*. Preprint at <http://arxiv.org/abs/1904.09075>
- Amgad M, Elfandy H, Hussein H, Atteya LA, Elsebaie MAT, Elnasr LSA, Sakr RA, Salem HSE, Ismail AF, Saad AM et al (2019) Structured crowdsourcing enables convolutional segmentation of histology images. *Bioinformatics* 35(18):3461–3467
- Arganda-Carreras I et al (2015) Crowdsourcing the creation of image segmentation algorithms for connectomics. *Front Neuroanat* 9:142
- Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, Zhao B, Aberle DR, Henschke CI, Hoffman EA et al (2011) The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Med Phys* 38(2):915–931 (**PubMed: 21452728**)
- Basha SS, Ghosh S, Babu KK, Dubey SR, Pulabaigari V, Mukherjee S (2018) Rccnet: an efficient convolutional neural network for histological routine colon cancer nuclei classification. In: 2018 15th international conference on control, automation, robotics and vision (ICARCV). *IEEE*, pp 1222–1227
- Bernal J, Sánchez FJ, Fernández-Esparrach G, Gil D, Rodríguez C, Vilariño F (2015) WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics* 43:99–111
- Bernal J, Tajbaksh N, Sánchez FJ, Matuszewski BJ, Chen H, Yu L, Angermann Q, Romain O, Rustad B, Balasingham I et al (2017) Comparative validation of polyp detection methods in video colonoscopy: results from the miccai 2015 endoscopic vision challenge. *IEEE Trans Med Imaging* 36(6):1231–1249
- Birodkar V, Lu Z, Li S, Rathod V, Huang J (2021) The surprising impact of mask-head architecture on novel class segmentation. Preprint at [arXiv:2104.00613](https://arxiv.org/abs/2104.00613)
- Buda M (2020) Brain mri segmentation. [Online]. <https://www.kaggle.com/mateuszbeda/igg-mri-segmentation>
- Buda M, Saha A, Mazurowski MA (2019) Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. *Comput Biol Med* 109:218–225
- Budginaite E, Morkūnas M, Laurinavičius A, Treigys P (2021) Deep learning model for cell nuclei segmentation and lymphocyte identification in whole slide histology images. *Informatika* 32(1):23–40
- Caicedo JC et al (2019) Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat Methods* 16(12):1247–1253
- Camalan S, Mahmood H, Binol H, Araújo ALD, Santos-Silva AR, Vargas PA et al (2021) Convolutional neural network-based clinical predictors of oral dysplasia: class activation map analysis of deep learning results. *Cancers*. <https://doi.org/10.3390/cancers13061291>
- Candemir S, Jaeger S, Palaniappan K, Musco JP, Singh RK, Xue Z, Karargyris A, Antani S, Thoma G, McDonald CJ (2013) Lung segmentation in chest radiographs using anatomical atlases with non-rigid registration. *IEEE Trans Med Imaging* 33:577–590
- Cardona A et al (2010) An integrated micro- and macroarchitectural analysis of the Drosophila brain by computer-assisted serial section electron microscopy. *PLoS Biol* 8:e1000502
- Celik Y, Talo M, Yildirim O, Karabatak M, Acharya UR (2020) Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. *Pattern Recogn Lett* 133:232–239
- Cervantes-Sanchez F, Maktabi M, Köhler H, Sucher R, Rayes N, Avina-Cervantes JG et al (2021) Automatic tissue segmentation of hyperspectral images in liver and head neck surgeries using machine learning. *Artif Intell Surg* 1:22–37
- Chanchal AK, Lal S, Kini J (2021a) High resolution deep transferred ASPPU-net for nuclei segmentation of histopathology images. *Int J Comput Assist Radiol Surg*. <https://doi.org/10.1007/s11548-021-02497-9>. (**PMID: 34622381**)
- Chanchal AK, Kumar A, Lal S, Kini J (2021b) Efficient and robust deep learning architecture for segmentation of kidney and breast histopathology images. *Comput Electr Eng* 92:107177
- Chen L, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2018) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans Pattern Anal Mach Intell* 40(4):834–848
- Chen S, Ding C, Tao D (2020) Boundary-assisted region proposal networks for nucleus segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, pp 279–288
- Chen S, Ding C, Liu M, Tao D (2021) CPP-Net: Context-aware polygon proposal network for nucleus segmentation. Preprint at [arXiv:2102.06867](https://arxiv.org/abs/2102.06867)
- Chidester B, Ton T-V, Tran M-T, Ma J, Do MN (2019) Enhanced rotation-equivariant U-net for nuclear segmentation. In: *Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*, Long Beach, 16–17 June 2019, pp 1097–1104
- Cicconet M, Hochbaum DR, Richmond DL, Sabatin BL (2017) Bots for software-assisted analysis of image-based transcriptomics. In: *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, pp 134–142
- Codella NC, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW et al (2018) Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). *IEEE*, pp 168–172



- Codella NCF, Rotemberg V, Tschandi P, Celebi ME, Dusza SW, Gutman D, Helba B, Kalloo A, Liopyris K, Marchetti MA, Kittler H, Halpern A (2019) Skin lesion analysis toward melanoma detection 2018a: A challenge hosted by the international skin imaging collaboration (ISIC), CoRRabs/1902.03368
- Cohen JP, Morrison P, Dao L (2020) Covid-19 image data collection. Preprint at [arXiv:2003.11597](https://arxiv.org/abs/2003.11597)
- Cruz-Roa A, Basavanthally A, González F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A (2014) Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Med. Imaging 2014 Digit. Pathol. SPIE, pp 904103. <https://doi.org/10.1117/12.2043872>.
- Data science bowl (2018) <https://www.kaggle.com/c/data-science-bowl-2018>
- Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) ImageNet: a large-scale hierarchical image database. In: CVPR09
- Dewan MAA, Ahmad MO, Swamy MNS (2011) Tracking biological cells in time-lapse microscopy: an adaptive technique combining motion and topological features. *IEEE Trans Biomed Eng* 58(6):1637–1647
- Dinh TL, Kwon SG, Lee SH, Kwon KR (2021) Breast tumor cell nuclei segmentation in histopathology images using EfficientUnet++ and multi-organ transfer learning. *J Korea Multimed Soc* 24(8):1000–1011
- Dogan RO, Dogan H, Bayrak C, Kayikcioglu T (2021) A two-phase approach using mask R-CNN and 3D U-net for high-accuracy automatic segmentation of pancreas in CT imaging. *Comput Methods Programs Biomed* 207:106141
- Elmore JG et al (2015) Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 313:1122–1132
- Ethical approval for the Sheffield cohort was obtained for this study from the HRA and Health and Care Research Wales (HCRW), Reference number 18/WM/0335 on 19 October 2018
- Feng L, Song JH, Kim J, Jeong S, Park JS, Kim J (2019) Robust nucleus detection with partially labeled exemplars. *IEEE Access* 7:162169–162178
- Feng Y, Hafiane A, Laurent H (2020) A deep learning based multiscale approach to segment cancer area in liver whole slide image. Preprint at [arXiv:2007.12935](https://arxiv.org/abs/2007.12935)
- Fishman D, Salumaa S-O, Majoral D et al (2019) Segmenting Nuclei in Brightfield Images with Neural Networks. *bioRxiv*. <https://doi.org/10.1101/764894>
- Gamper J, Koohbanani NA, Benet K, Khuram A, Rajpoot N (2019) PanNuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In: Proc. Eur. Congr. Digit. Pathol. (ECDP), pp 11–19
- Gamper J et al (2020) PanNuke dataset extension, insights and baselines. Preprint at [arXiv:2003.10778](https://arxiv.org/abs/2003.10778)
- Gong X, Chen S, Zhang B, Doermann D (2021) Style consistent image generation for nuclei instance segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 3994–4003
- Graham S, Vu QD, Raza SEA, Azam A, Tsang YW, Kwak JT, Rajpoot N (2019) Hover-net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med Image Anal* 58:101563
- Grishagin IV (2015) Automatic cell counting with ImageJ. *Anal Biochem* 473:63–65
- Gudhe NR, Behravan H, Sudah M, Okuma H, Vanninen R, Kosma VM, Mannermaa A (2021) Multi-level dilated residual network for biomedical image segmentation. *Sci Rep* 11(1):1–18
- Han JW, Breckon TP, Randell DA, Landini G (2008) Radicular cysts and odontogenic keratocysts epithelia classification using cascaded Haar classifiers (PDF). In: Proc. 12th annual conference on medical image understanding and analysis, pp 54–58 (Retrieved 8 April 2013)
- Han JW, Breckon TP, Randell DA, Landini G (2012) The application of support vector machine classification to detect cell nuclei for automated microscopy. *Mach vis Appl* 23(1):15–24. <https://doi.org/10.1007/s00138-010-0275-y>. (Retrieved 8 April 2013)
- Hassan L, Saleh A, Abdel-Nasser M, Omer OA, Puig D (2021a) Promising deep semantic nuclei segmentation models for multi-institutional histopathology images of different organs. *Int J Interact Multimed Artif Intell* 6(6)
- Hassan L, Saleh A, Abdel-Nasser M, Omer OA, Puig D (2021b) Efficient multi-organ multi-center cell nuclei segmentation method based on deep learnable aggregation network. *Traitement Du Signal* 38(3):653–661
- Hayakawa T, Prasath VB, Kawanaka H, Aronow BJ, Tsuruoka S (2021) Computational nuclei segmentation methods in digital pathology: a survey. *Arch Comput Methods Eng* 28(1):1–13
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. *IEEE Conf Comput vis Pattern Recog (CVPR)* 2016:770–778
- Ioannidis GS, Trivizakis E, Metzakis I, Papagiannakis S, Lagoudaki E, Marias K (2021) Pathomics and deep learning classification of a heterogeneous fluorescence histology image dataset. *Appl Sci* 11(9):3796
- Irshad H, Veillard A, Roux L, Racoceanu D (2013) Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE Rev Biomed Eng* 7:97–114
- Irshad H, Kouhsari LM, Waltz G, Bucur O, Nowak JA, Dong F, Knoblauch NW, Beck AH (2015) Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd. In: Pacific symposium on biocomputing (PSB). pp 294–305. <https://doi.org/10.13140/2.1.4067.0721>
- Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Xue Z, Palaniappan K, Singh RK, Antani S et al (2013) Automatic tuberculosis screening using chest radiographs. *IEEE Trans Med Imaging* 33:233–245
- Jahanifar M, Tajeddin NZ, Koohbanani NA, Rajpoot N (2021) Robust interactive semantic segmentation of pathology images with minimal user input. Preprint at [arXiv:2108.13368](https://arxiv.org/abs/2108.13368)
- Janowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J Pathol Inform*. <https://doi.org/10.4103/2153-3539.186902>
- Jantzen J, Norup J, Dounias G, Bjerregaard B (2005) Pap-smear benchmark data for pattern classification. *Nature inspired Smart Information Systems (NiSIS 2005)*, pp 1–9
- Jevtic P, Edens LJ, Vukovic LD, Levy DL (2014) Sizing and shaping the nucleus: mechanisms and significance. *Curr Opin Cell Biol* 28:16–27. <https://doi.org/10.1016/j.ceb.2014.01.003>
- Jha D, Riegler MA, Johansen D, Halvorsen P, Johansen HD (2020) Doubleu-net: a deep convolutional neural network for medical image segmentation. In: 2020 IEEE 33rd International symposium on computer-based medical systems (CBMS). IEEE, pp 558–564
- 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
- Kadia DD, Alom MZ, Burada R, Nguyen TV, Asari VK (2021) R2U3D: recurrent residual 3D U-net for lung segmentation. Preprint at [arXiv:2105.02290](https://arxiv.org/abs/2105.02290)
- Kanadath A, Jothi JAA, Urolagin S (2021) Histopathology image segmentation using MobileNetV2 based U-net model. In: 2021 international conference on intelligent technologies (CONIT). IEEE, pp 1–8



- Kang Q, Lao Q, Fevens T (2019) Nuclei segmentation in histopathological images using two-stage learning. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 703–711
- Kasturba Medical College (KMC) (2021) Mangalore, Manipal Academy of Higher Education (MAHE), Manipal, Karnataka, India for sharing liver cancer histopathology image dataset
- Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA et al (2019) Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. *PLoS Med* 16:e1002730
- Ke J, Shen Y, Lu Y, Deng J, Wright JD, Zhang Y et al (2021) Quantitative analysis of abnormalities in gynecologic cytopathology with deep learning. *Lab Invest* 101(4):513–524
- Ker J, Wang L, Rao J, Lim T (2017) Deep learning applications in medical image analysis. *IEEE Access* 6:9375–9389
- Khan AR, Khan S, Harouni M, Abbasi R, Iqbal S, Mehmood Z (2021) Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. *Microscopy Research and Technique*
- Kimura H, Yonemura Y (1991) Flow cytometric analysis of nuclear DNA content in advanced gastric cancer and its relationship with prognosis. *Cancer* 67(10):2588–2593
- Kong Y, Genchev GZ, Wang X, Zhao H, Lu H (2020) Nuclear segmentation in histopathological images using two-stage stacked U-nets with attention mechanism. *Front Bioeng Biotechnol* 8:1246
- Koohbanani NA, Jahanifar M, Gooya A, Rajpoot N (2019) Nuclear instance segmentation using a proposal-free spatially aware deep learning framework. In: International conference on medical image computing and computer-assisted intervention. Springer, Cham, pp 622–630
- Kowal M, Filipczuk P (2014) Nuclei segmentation for computer-aided diagnosis of breast cancer. *Int J Appl Math Comput Sci* 24(1):19–31
- Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90
- Kromp F, Bozsaky E, Rifatbegovic F, Fischer L, Ambros M, Berneder M, Weiss T, Lazić D, Dörr W, Hanbury A et al (2020) An annotated fluorescence image dataset for training nuclear segmentation methods. *Sci Data* 7:1–8
- Kumar N, Verma R, Sharma S, Bhargava S, Vahadane A, Sethi A (2017) A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans Med Imaging* 36(7):1550–1560
- Kumar N et al (2020) A multi-organ nucleus segmentation challenge. *IEEE Trans Med Imaging*. <https://doi.org/10.1109/TMI.2019.2947628>
- Lagree A, Mohebpour M, Meti N, Saednia K, Lu FI, Slodkowska E et al (2021) A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks. *Sci Rep* 11(1):1–11
- Lal S, Das D, Alabhya K, Kanfode A, Kumar A, Kini J (2021) NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images. *Comput Biol Med* 128:104075
- Li W (2015) Automatic segmentation of liver tumor in ct images with deep convolutional neural networks. *J Comput Commun* 3(11):146–151
- Li J, Hu Z, Yang S (2019a) Accurate nuclear segmentation with center vector encoding. In: International conference on information processing in medical imaging. Springer, Cham, pp 394–404
- Li C et al (2019b) Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Med Image Anal* 53:165–178
- Li L, Wei M, Liu B, Atchaneeyasakul K, Zhou F, Pan Z, Kumar SA, Zhang JY, Pu Y, Liebeskind DS, Scalzo F (2020) Deep learning for hemorrhagic lesion detection and segmentation on brain CT images. *IEEE J Biomed Health Inform* 25(5):1646–1659
- Li Y, Wu X, Li C, Sun C, Li X, Rahaman M, Zhang Y (2021a) Intelligent gastric histopathology image classification using hierarchical conditional random field based attention mechanism. In: 2021a 13th international conference on machine learning and computing, pp 330–335
- Li X, Yang H, He J, Jha A, Fogo AB, Wheless LE et al (2021b) BEDS: bagging ensemble deep segmentation for nucleus segmentation with testing stage stain augmentation. In: 2021b IEEE 18th international symposium on biomedical imaging (ISBI). IEEE, pp 659–662
- Lin TY, Maire M, Belongie S, Bourdev L, Girshick R, Hays J, Perona P, Ramanan D, Zitnick CL, Dollar P (2015) Microsoft COCO: Common Objects in Context. Preprint at [arXiv:1405.0312](https://arxiv.org/pdf/1405.0312). [Online]. <https://arxiv.org/pdf/1405.0312>
- Liu J, Xu B, Zheng C, Gong Y, Garibaldi J, Soria D et al (2018a) An end-to-end deep learning histochemical scoring system for breast cancer TMA. *IEEE Trans Med Imaging* 38(2):617–628
- Liu Y, Zhang P, Song Q, Li A, Zhang P, Gui Z (2018b) Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. *IEEE Access* 6:53709–53721
- Liu D, Zhang D, Song Y, Zhang C, Zhang F, O'Donnell L, Cai W (2019) Nuclei segmentation via a deep panoptic model with semantic feature fusion. *IJCAI*, pp 861–868
- Liu X, Guo Z, Cao J, Tang J (2021a) MDC-Net: a new convolutional neural network for nucleus segmentation in histopathology images with distance maps and contour information. *Comput Biol Med* 135:104543
- Liu K, Mokhtari M, Li B, Nofallah S, May C, Chang O et al (2021b) Learning melanocytic proliferation segmentation in histopathology images from imperfect annotations. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 3766–3775
- Ljosa V, Sokolnicki KL, Carpenter AE (2012) Annotated high-throughput microscopy image sets for validation. *Nat Methods* 9(7):637–637
- Louis DN et al (2015) Computational pathology: a path ahead. *Arch Pathol Lab Med* 140(1):41–50
- Lu C, Romo-Bucheli D, Wang X, Janowczyk A, Ganesan S, Gilmore H, Rimm D, Madabhushi A (2018) Nuclear shape and orientation features from h&e images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab Invest* 98(11):1438
- LUNA16—Home (2020) [Online]. <https://luna16.grand-challenge.org/>. Accessed 4 Nov 2020
- Mahbod A, Schaefer G, Ellinger I, Ecker R, Smedby Ö, Wang C (2019) A two-stage U-Net algorithm for segmentation of nuclei in H&E-stained tissues. In: European congress on digital pathology. Springer, Cham, pp 75–82
- Mahmood T, Arsalan M, Owais M, Lee MB, Park KR (2020) Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster R-CNN and Deep CNNs. *J Clin Med* 9:749
- Mahmood T, Owais M, Noh KJ, Yoon HS, Koo JH, Haider A et al (2021) Accurate segmentation of nuclear regions with multi-organ histopathology images using artificial intelligence for cancer diagnosis in personalized medicine. *J Personal Med* 11(6):515
- Maktabi M, Köhler H, Ivanova M et al (2020) Classification of hyperspectral endocrine tissue images using support vector machines. *Int J Med Robot* 16:1–10
- Mehta S, Lu X, Weaver D, Elmore JG, Hajishirzi H, Shapiro L (2020) HATNet: an end-to-end holistic attention network for diagnosis of breast biopsy images. Preprint at [arXiv:2007.13007](https://arxiv.org/abs/2007.13007)

- Meijering E, Dzyubachyk O, Smal I, van Cappellen WA (2009) Tracking in cell and developmental biology. *Semin Cell Dev Biol* 20(8):894–902
- Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J et al (2014) The multimodal brain tumor image segmentation benchmark (BraTS). *IEEE Trans Med Imaging* 34(10):1993–2024
- Natarajan VA, Kumar MS, Patan R, Kallam S, Mohamed MYN (2020) Segmentation of nuclei in histopathology images using fully convolutional deep neural architecture. In: 2020 International conference on computing and information technology (ICCIIT-1441). IEEE, pp 1–7
- Naylor P, Laé M, Reyat F, Walter T (2017) Nuclei segmentation in histopathology images using deep neural networks. In: *Bio-medical imaging (ISBI 2017)*, 2017 IEEE 14th international symposium on. IEEE, pp 933–936. <https://doi.org/10.1109/isbi.2017.7950669>
- Naylor P, La'e M, Reyat F, Walter T (2018) Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans Med Imaging* 38(2):448–459. <https://doi.org/10.1109/TMI.2018.2865709>
- Özyurt F, Sert E, Avci E, Dogantekin E (2019) Brain tumor detection based on convolutional neural network with neutrosophic expert maximum fuzzy sure entropy. *Measurement* 147:106830
- Paeng K, Hwang S, Park S, Kim M (2017) A Unified framework for tumor proliferation score prediction in breast histopathology. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10553 LNCS, pp 231–239
- PAIP2019 (2019) <https://paip2019.grand-challenge.org/>
- Plissiti ME, Dimitrakopoulos P, Sfikas G, Nikou C, Krikoni O, Charchanti A (2018) SIPAKMED: a new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images. In: 2018 25th IEEE international conference on image processing (ICIP). IEEE, pp 3144–3148
- Piracicaba Dental Ethical Committee (2019) Registration number 42235421.9.0000.5418
- Podder S, Bhattacharjee S, Roy A (2021) An efficient method of detection of COVID-19 using Mask R-CNN on chest X-Ray images. *AIMS Biophysics* 8(3):281–290
- Porzi L, Buló SR, Penate-Sanchez A, Ricci E, Moreno-Noguer F (2016) Learning depth-aware deep representations for robotic perception. *IEEE Robot Autom Lett* 2(2):468–475
- Qu H, Wu P, Huang Q, Yi J, Yan Z, Li K et al (2020) Weakly supervised deep nuclei segmentation using partial points annotation in histopathology images. *IEEE Trans Med Imaging* 39(11):3655–3666
- Rabbani M (2002) JPEG2000: Image compression fundamentals, standards and practice. *J Electron Imaging* 11(2):286
- Rai R, Das A, Dhal KG (2022) Nature-inspired optimization algorithms and their significance in multi-thresholding image segmentation: an inclusive review. *Evol Syst*. <https://doi.org/10.1007/s12530-022-09425-5>
- Reza MS, Ma J (2018) Imbalanced histopathological breast cancer image classification with convolutional neural network. In: 14th IEEE international conference on signal processing (ICSP), pp 619–624
- Romero FP, Tang A, Kadoury S (2019) Multi-level batch normalization. In: *Deep networks for invasive ductal carcinoma cell discrimination in histopathology images*. Preprint at [arXiv:1901.03684](https://arxiv.org/abs/1901.03684)
- Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, pp 234–241
- Roth HR, Lu L, Lay N, Harrison AP, Farag A, Sohn A, Summers RM (2018) Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation. *Med Image Anal* 45:94–107
- Roy M, Kong J, Kashyap S, Pastore VP, Wang F, Wong KC, Mukherjee V (2021) Convolutional autoencoder based model histocae for segmentation of viable tumor regions in liver whole-slide images. *Sci Rep* 11(1):1–10
- Schols RM, terLaan M, Stassen LP et al (2014) Differentiation between nerve and adipose tissue using wide-band (350–1,830 nm) in vivo diffuse reflectance spectroscopy. *Lasers Surg Med* 46:538–545
- Schols RM, Alic L, Wieringa FP, Bouvy ND, Stassen LP (2017) Towards automated spectroscopic tissue classification in thyroid and parathyroid surgery. *Int J Med Robot* 13:e1748
- Seetha J, Raja SS (2018) Brain tumor classification using convolutional neural networks. *Biomed Pharmacol J* 11:1457–1461
- Shuvo MB, Ahommed R, Reza S, Hashem MMA (2021) CNL-UNet: a novel lightweight deep learning architecture for multimodal biomedical image segmentation with false output suppression. *Biomed Signal Process Control* 70:102959
- Silva AB, Martins AS, Neves LA, Faria PR, Tosta TA, do Nascimento MZ (2019) Automated nuclei segmentation in dysplastic histopathological oral tissues using deep neural networks. In: *Ibero-american congress on pattern recognition*. Springer, Cham, pp 365–374
- Sirinukunwattana K, Raza SEA, Tsang Y-W, Snead DR, Cree IA, Rajpoot NM (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
- Sohail A, Khan A, Wahab N, Zameer A, Khan S (2021) A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images. *Sci Rep* 11(1):1–18
- Song T-H, Sanchez V, EIDaly H, Rajpoot NM (2017) Dual-channel active contour model for megakaryocytic cell segmentation in bone marrow trephine histology images. *IEEE Trans Bio-Med Eng* 64(12):2913–2923
- Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2015) A dataset for breast cancer histopathological image classification. *IEEE Trans Biomed Eng* 63:1455–1462
- Su H, Xing F, Kong X, Xie Y, Zhang S, Yang L (2015) Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, pp 383–390
- Szeliski R (2010) *Computer vision: algorithms and applications*. Springer Science & Business Media, Berlin
- Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging* 35(5):1299–1312 (**PubMed**: 26978662)
- Tarighat AP (2021) Breast tumor segmentation using deep learning by U-net network. *J Telecommun Electron Comput Eng (JTEC)* 13(2):49–54
- The Cancer Genome Atlas (TCGA) (2016) [Online]. <http://cancergenome.nih.gov/>. Accessed 14 May 2016
- Toğaçar M, Özkurt KB, Ergen B, Cömert Z (2020) Breastnet: a novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. *Phys A Stat Mech Appl* 545:123592. <http://www.sciencedirect.com/science/article/pii/S0378437119319995>
- Tomczak K, Czerwińska P, Wiznerowicz M (2015) Review the cancer genome atlas (TCGA): an immeasurable source of knowledge. *WspolczesnaOnkol Oncol* 2015:68–77. <https://doi.org/10.5114/wo.2014.47136>

- Tschandl P, Rosendahl C, Kittler H (2018) The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 5:180161
- Ultrasound nerve segmentation (2016) <https://www.kaggle.com/c/ultrasound-nerve-segmentation>
- Vahadane A et al (2016) Structure-preserving color normalization and sparse stain separation for histological images. *IEEE Trans Med Imaging* 35:1962–1971
- VESSEL12—Home (2020) [Online]. <https://vessel12.grand-challenge.org/>. Accessed 4 Nov 2020
- Vivanti R, Ephrat A, Joskowicz L, Karaaslan O, Lev-Cohain N, Sosna J (2015) Automatic liver tumor segmentation in follow-up ct studies using convolutional neural networks. In: *Proc. Methods Med. Image Process. Workshop*, vol 2
- Vu QD, Graham S, To MNN, Shaban M, Qaiser T, Koohbanani NA, Khurram SA, Kurc T, Farahani K, Zhao T et al (2018) Methods for segmentation and classification of digital microscopy tissue images. Preprint at [arXiv:1810.13230](https://arxiv.org/abs/1810.13230)
- Wahab N, Khan A, Lee YS (2019) Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images. *Microscopy* 68:216–233
- Wang EK, Zhang X, Pan L, Cheng C, Dimitrakopoulou-Strauss A, Li Y, Zhe N (2019a) Multi-path dilated residual network for nuclei segmentation and detection. *Cells* 8(5):499
- Wang S, Zhu Y, Yu L, Chen H, Lin H, Wan X et al (2019b) RMDL: recalibrated multi-instance deep learning for whole slide gastric image classification. *Med Image Anal* 58:101549
- Wang H, Xian M, Vakanski A (2020) Bending loss regularized network for nuclei segmentation in histopathology images. In: *2020 IEEE 17th international symposium on biomedical imaging (ISBI)*. IEEE, pp 1–5
- Wang H, Vakanski A, Shi C, Xian M (2021) Bend-Net: bending loss regularized multitask learning network for nuclei segmentation in histopathology images. Preprint at [arXiv:2109.15283](https://arxiv.org/abs/2109.15283)
- Wenzhong L, Huanlan L, Caijian H, Liangjun Z (2020) Classifications of breast cancer images by deep learning. *medRxiv*
- Woo S, Park J, Lee J-Y, Kweon IS (2018) Cbam: convolutional block attention module. *Lecture Notes Comput Sci*. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)
- Xiao W, Jiang Y, Yao Z, Zhou X, Lian J, Zheng Y (2021) Polar representation-based cell nucleus segmentation in non-small cell lung cancer histopathological images. *Biomed Signal Process Control* 70:103028
- Yoo I, Yoo D, Paeng K (2019) Pseudoedgenet: nuclei segmentation only with point annotations. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, pp 731–739
- Yu J-M, Yang L-H et al (1989) Flow cytometric analysis DNA content in esophageal carcinoma: correlation with histologic and clinical features. *Cunccir* 64:80–82
- Zeng Z, Xie W, Zhang Y, Lu Y (2019) RIC-Unet: an improved neural network based on Unet for nuclei segmentation in histology images. *IEEE Access* 7:21420–21428
- Zhang Z, Lin C (2018) Pathological image classification of gastric cancer based on depth learning. *ACM Trans Intell Syst Technol* 45(11A):263–268
- Zhao J, Li Q, Li X, Li H, Zhang L (2019a) Automated segmentation of cervical nuclei in pap smear images using deformable multi-path ensemble model. In: *2019a IEEE 16th international symposium on biomedical imaging (ISBI 2019a)*. IEEE, pp 1514–1518
- Zhao J, Dai L, Zhang M, Yu F, Li M, Li H et al (2019b) PGU-net+: progressive growing of U-net+ for automated cervical nuclei segmentation. In: *International workshop on multiscale multi-modal medical imaging*, pp 51–58
- Zhou Y, Xie L, Shen W, Fishman E, Yuille (2016) A Pancreas segmentation in abdominal ct scan: a coarse-to-fine approach. Preprint at [arXiv:1612.08230](https://arxiv.org/abs/1612.08230).
- Zhou Z, Shin J, Zhang L, Gurudu S, Gotway M, Liang J (2017) Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally. In: *IEEE conference on computer vision and pattern recognition (CVPR)*, pp 7340–7351
- Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J (2018) Unet++: a nested u-net architecture for medical image segmentation. In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, Cham, pp 3–11
- Zhou Y, Onder OF, Dou Q, Tsougenis E, Chen H, Heng PA (2019a) Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In: *International conference on information processing in medical imaging*. Springer, Cham, pp 682–693
- Zhou Y, Chen H, Xu J, Dou Q, Heng PA (2019b) Irnet: instance relation network for overlapping cervical cell segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, pp 640–648
- Zhou Y, Chen H, Lin H, Heng PA (2020) Deep semi-supervised knowledge distillation for overlapping cervical cell instance segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, pp 521–531
- Zhou C, Jin Y, Chen Y, Huang S, Huang R, Wang Y et al (2021) Histopathology classification and localization of colorectal cancer using global labels by weakly supervised deep learning. *Comput Med Imaging Graph* 88:101861
- Zunair H, Hamza AB (2021) Sharp U-Net: depthwise convolutional network for biomedical image segmentation. *Comput Biol Med* 136:104699

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.