Multimedia Tools and Applications
https://doi.org/10.1007/s11042-022-14098-y

TRACK 2: MEDICAL APPLICATIONS OF MULTIMEDIA

AlexSegNet: an accurate nuclei segmentation deep learning model in microscopic images for diagnosis of cancer

Anu Singha [1] · Mrinal Kanti Bhowmik [2]

Received: 22 May 2022 / Revised: 10 August 2022 / Accepted: 21 October 2022

Abstract
The nuclei segmentation of microscopic images is a key pre-requisite for cancerous pathological image analysis. However, an accurate nuclei cell segmentation is a long running major challenge due to the enormous color variability of staining, nuclei shapes, sizes, and clustering of overlapping cells. To address this challenges, we proposed a deep learning model, namely, AlexSegNet which is based upon AlexNet model Encoder-Decoder framework. In Encoder part, it stitches feature maps in the channel dimension to achieve feature fusion and uses a skip structure in Decoder part to combine low- and high-level features to ensure the segmentation effect of the nucleus. At final stage, we have also introduced a stacked network where feature maps are stacks on top of each other. We have used a publically available 2018 Data Science Bowl and Triple Negative Breast Cancer (TNBC) datasets of microscopic nuclei images for this study which comprises of several sample types such as small and large fluorescent, pink, purple, and grayscale tissue samples. Experimental results show that our proposed AlexSegNet achieved a segmentation maximum performance of 91.66% for Data Science Bowl dataset and 66.88% for TNBC dataset. The results are competitive compared to the results of other state-of-the-art models. This model is expected to be useful clinically for technician experts to succeed the analysis of cancer diagnosis into the survival chances of patients.

✉ Anu Singha
anusingh5012@gmail.com

Mrinal Kanti Bhowmik
mrinalkantibhowmik@tripurauniv.ac.in

[1] Department of Computer Science and Engineering, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad 201204 Uttar Pradesh, India

[2] Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar-799022, Agartala, India

Published online: 01 December 2022

Keywords Convolutional neural network · Nuclei · Segmentation · Fluorescent · Histopathology · cancer

1 Introduction

The segmentation of nuclei from microscopic images has persisted the cancer research community's key focus for decades. In human's health concerns, cancer diseases become the most common and life-threatening issue. In India, the incidence of various cancer growth rate increases in the young age group at very aggressive [45]. Cancer starts from a benign state and at the early stages without proper treatment, it becomes malignant when the nuclei cells start to grow abnormally. According to international agency for research on cancer, world health organization (WHO), the worldwide estimated number of new cases in 2020, the breast cancer is the highest possible with 2,261,419 (11.7%), as shown in Fig. 1. The availability of appropriate screening devices is essential for detecting the initial symptoms of cancer. Numerous imaging techniques are used for the screening to detect this disease such as mammography, thermography, histopathology, and dermoscopy [5, 6]. To improve the accuracy of the diagnosis for patients, the histopathology microscopic images are considered as the gold standard among other imaging techniques [6]. Moreover, the histopathological examination can deliver more inclusive and reliable evidence to diagnose cancer and measure its effects on the surrounding tissues [3–9].

For analysis of microscopic histopathology images, the detection and segmentation of nuclei cells are essential steps. These segmented nuclei are used in the grading diagnosis of many cancers which require comprehensive analysis of the characteristics of the nuclei such as shape, size, gray value, color variation of samples, clusters of nuclei with overlapping, and ratio of nuclei to cytoplasm. It is a major challenge in microscopic histopathology images of different patients where the shape and appearance of the different nuclei for disease stages vary greatly. Suppose in breast cancer, the identification of the stages of aggressiveness of the

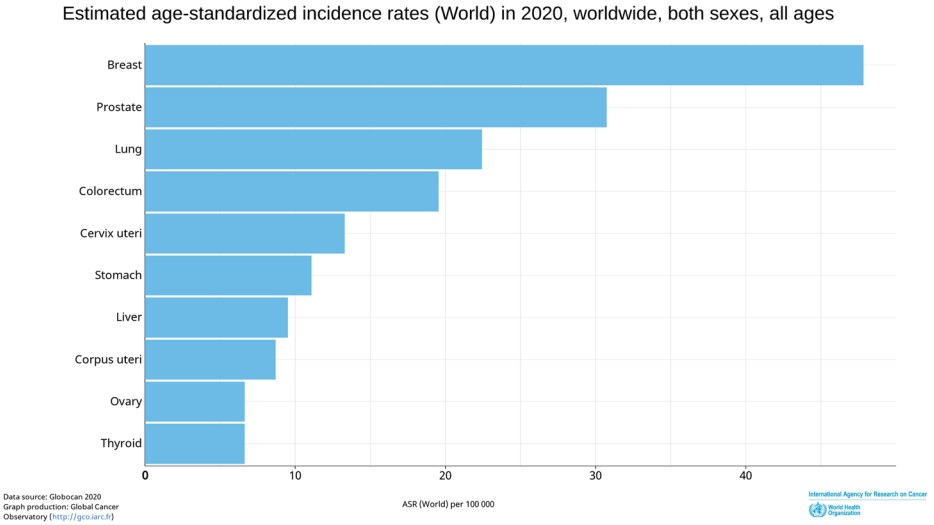


Fig. 1 Estimated number of new cases in 2020, worldwide, both sexes, all ages [20]

disease based upon the Nottingham Histologic Score system which also largely based off the morphologic attributes of the histopathology nuclei [7]. As a consequence, the accurate segmentation of histopathology image nuclei is a great challenging work in developing automated machine via computer assisted decision support for digital histopathology. Other than these, the segmentation of nuclei of histopathology images is essential to numerous studies, such as feature extraction, cell counts, and classification.

There are several approaches to handle segmentation of nuclei. Classical approach for cell image segmentation methods include region growing algorithm, threshold segmentation, edge detection, morphology and classification-clustering methods [12–14]. Another approach very suitable for analysis variability and complexity of medical images is deep learning. Deep convolutional network have ability to learn and coarse representations layer by layer process pathology images of several nuclei parts. Recently, the models include SegNet [4], ENet [41], PSPNet [51], U-Net [44], ESPNet [30], SPNetv2 [32], ICNet [52] and Y-Net [31]. They have achieved reasonably great performance on the segmentation compared to traditional medical image segmentation methods: to resolve the problems of poor segmentation of small nucleus, and under- and over-segmentation.

Even though convolutional models achieved impressive results, there is some flaws. The boundaries of segmented image by ENet are relatively blurred, due to insufficient deployment of shallow layer structure information. SegNet network is not well integrated the high-level semantic information with the shallow image information which produces segmented image with more noise points and blurred edges. Due to the limited size of receptive field in U-Net, the perception ability of cluster nucleus is weak which leads to under-segmentation. To overcome the above difficulties and challenges, this paper introduces a new segmentation network, namely, AlexSegNet.

The contributions of this paper as follows:

i. AlexSegNet model is proposed to handle nucleus segmentation of histopathology images for cancer treatment. This model is based upon AlexNet model Encoder-Decoder framework. In Encoder part, it stitches feature maps in the channel dimension to achieve feature fusion and uses a skip structure in Decoder part to combine low- and high-level features to ensure the segmentation effect of the nucleus.
ii. Study the performance of proposed model with nuclei segmentation in microscopic different colored histopathology, fluorescent nuclei images from publicly available datasets. It achieves sufficient improvement into the measured metrics.
iii. A comprehensive comparative study with existing state-of-the-arts and related methods is presented, in terms of evaluation via recall, precision, and F_1-score metrics.

The rest of paper is organized as follows: Section 2 describes the related work, and Section 3 describes elaborately the proposed deep learning framework in this study. Section 4 the experimental results, discussions, and comparison with the state-of-the-art. Section 5 estimates the computational complexity. Finally, the paper is concluded by Section 6.

2 Related work

Over the recent times, convolutional neural network based analysis of several medical image modalities has been demanded highly due to its robust performance. Now, we will present

overview of the important recent developments in the medical image segmentation including nuclei of histopathology images. Since most disease analysis highly dependent on cell-level information, the segmentation challenge in computing technology is to investigate all the individual cells to make correct diagnosis. Table 1 shows briefly the pros and cons for each listed related works in this section.

A fully convolutional network (FCN) [28] was the initial full convolution neural network which has been introduced for medical image segmentation. The authors defined a skip

Table 1 List of related works pros and cons briefly

Literature Works	Pros	Cons
J. Long et al. [28]	Define a skip architecture to take advantage of the feature spectrum that combines deep, coarse, semantic information and shallow, fine, appearance information	–
Y. Cui et al. [10]	The idea of simultaneous nucleus-boundary identification approach which can be applied to other biomedical image segmentation tasks such as gland segmentation and bacteria segmentation.	There is need to do some cropping operation in neural network training to make the size of layers match each other, which might lose useful surrounding information.
Ronneberger et al. [44]	It allows the network to learn invariance to deformations. This is particularly important in biomedical segmentation, since deformation used to be the most common variation in tissue and realistic deformations can be simulated efficiently.	Segmentation results in blurred effects.
S. Lal et al. [25]	Addressing shape-variability and nuclei-touching challenges during segmentation.	Computational complexity in terms of the number of parameters used and more convolutional layers.
L. Hassan et al. [17]	Nuclei segmentation in whole slide images of different stains and various organs.	–
H. Wang et al. [36]	Specially designed to handle overlap cells.	Lacking performance in different image scales.
S. Grahama et al. [46]	Accurate segmentation, particularly in areas with overlapping instances.	The model is trained on a single tissue type.
H. Su et al. [49]	Sparse reconstruction handles the shape variations of cells and touching cells in detection and segmentation.	–
J. Xu et al. [21]	Provide accurate seed points or vertices for developing cell-by-cell graph features.	–
A. Lagree et al. [23]	Accurately segment invasive carcinoma of the breast.	The study is on the relatively small dataset used for training and testing.
Kong Y et al. [12]	Segmentation of nuclei and overlapping regions.	–
Debesh J. et al. [1]	Model evaluated on various segmentation tasks using several imaging modalities such as colonoscopy, dermoscopy, and microscopy. And achieved high performances.	A limitation of the DoubleU-Net is that it uses more parameters as compared to U-Net, which leads to an increase in the training time.
Abhishek S. et al. [2]	Allows the preservation of resolution, improved information flow and propagation of both high- and low-level features to obtain accurate segmentation maps.	The model fails when extremely low contrast images are part of the data.
Ailiang L. et al. [50]	DS-TransUNet significantly outperforms especially in polyp segmentation task.	Not performing well in nuclei segmentation task

architecture to take benefit of feature information includes deep, coarse, semantic information and shallow, fine, appearance information. To produce fine segmentation, this network defined as a combination of deep semantic information with shallow appearance. Cui et al. [10] also introduced an automatic end-to-end fully convolutional neural network nucleus boundary model for segmenting of individual nuclei and their boundaries simultaneously. This simultaneous idea for nucleus-boundary identification can also be applied to other applications such as gland segmentation and bacteria segmentation. Ronneberger et al. [44] was the first who utilized U-Net as an encoder-decoder model for medical image segmentation tasks. U-Net uses skip connections to integrate low- and high-level information. This model allows to learn invariance to various deformations which particularly important in biomedical segmentation since deformation is the most common variation in tissue. Later-on, variety of U-Net extensions have been exploited for medical segmentation same purpose. H-DenseU-Net [26] for liver and tumor segmentation from CT scan, GP-U-Net [13] for lesion detection through segmentation tasks. These U-Net based models also achieve effective segmentation in histopathology images. A very popular semantic segmentation model SegNet [4] has been utilized by S. Lal et al. [25] and proposed an extension model NucleiSegNet to address the challenges of nuclei segmentation task in histopathology images of liver cancer. This model well-handled shape-variability and nuclei-touching tasks during segmentation. The NucleiSegNet architecture consists three blocks: a robust residual block, a bottleneck block, and an attention decoder block. Zhao et al. [51] proposed another semantic segmentation network, namely, PSPNet, inspired by context information of FCN. L. Hassan et al. [17] presented deep semantic nuclei segmentation model based upon PSPSegNet in histopathology images of different organs such as breast, kidney, prostate, and stomach.

For histopathology based nuclei segmentation task, Vu et al. [47] presented the DRAN model by integrating both nuclei and nuclei contours to achieve accurate nuclei segmentation results. CIA-Net [53] is another architecture for nuclei instance segmentation with contour-aware information aggregation. Instead of using independent decoders, this architecture exploits bi-directionally aggregated task-specific features instead of using independent decoders to model the texture and spatial dependencies between contour and nuclei. Histopathology nuclei segmentation based on a bending loss regularized network proposed by Wang et al. [36]. Minimizing bending loss can evade producing contours that comprise multiple nuclei. It uses the nuclei curvature to state high penalties for touching points of overlapped contour segments and assigns small penalties to well-stated nuclei contours. Naylor et al. [33] also presented a model 'DIST' for the segmentation of histopathology image nuclei by employing a regression concept for overlapping nuclei. Another deep learning work of Naylor et al. [16], combination of three models, namely PangNet, FCN, and DeconvNet, segmented the nuclei of triple negative breast cancer which then processed the posterior probability of the nuclei to accomplish the purpose of segmenting adhesion cells. Hover-Net [46] architecture was for simultaneous segmentation and classification of nuclei in histology images. Although this architecture was trained on a single tissue type, authors were confident that their architecture will achieve well if it incorporate additional tissue types. This claim based due to the robust performance of their instance segmentation architecture across multiple tissues. Recently, H. Jung et al. [26] proposed a nuclei segmentation model based on a backbone of Mask R-CNN for whole slide images (WSI).

Stacked Sparse auto-encoder is another successful approach for nuclei segmentation in histopathology samples. Auto-encoder is an encoder-decoder model where the encoder network characterises pixel intensities modeled via lower dimensional attributes, while the

decoder network rebuilds the original pixel intensities using these low dimensional features. Su et al. [49] proposed a stacked denoise autoencoder (SDAE) algorithm. This algorithm used for cell segmentation which deal with the challenges like gray inhomogeneity, shape changing, and cells overlapping. Xu et al. [21] also used stacked sparse autoencoder (SSAE) in breast cancer histopathology images but for detection of nuclei effectively. It provides precise seed points or vertices for evolving cell-by-cell graph features which can allow characterization of cellular topology features on tumor histology. More survey information about segmentation in histopathology nuclei images can be referred to H. Irshad et al. [48] and F. Xing et al. [24].

In this passage, we are listing few more recent works for accurate nuclei segmentation task. In 2021, Lagree et al. [23] studied a novel ensemble network, namely, GB U-Net to demonstrate the accurate breast invasive carcinoma tissue segmentation. A drawback of this study was the small dataset utilized for training and testing due to the limited open-source datasets. To solve the one of major challenge i.e. segmentation of nuclei as well as overlapping regions, Kong Y et al. [12] proposed a model based on two-stage stacked learning framework SUNets. First stage process the nuclei regions segmentation and the second stage divided the overlapping regions. Debesh J. et al. [1] improved the performance of U-Net by combining two U-Net networks stacked on top of each other and introduced a new model, namely, DoubleU-Net. The model evaluated on various segmentation tasks using several imaging modalities such as colonoscopy, dermoscopy, and microscopy. A limitation of the DoubleU-Net was that it uses additional parameters as compared to U-Net, which leads to a rise in the training time. Very recently, Abhishek S. et al. [2] proposed a novel network, namely, Multi-Scale Residual Fusion Network (MSRF-Net) for medical image segmentation which utilized a Dual-Scale Dense Fusion (DSDF) block to exchange multi-scale features of different receptive fields. The limitations of the proposed network is that it fails to extremely low contrast images for proper segmentation. Ailiang L. et al. [50] also introduced a model 'DS-TransUNet' for several medical image segmentation tasks. The authors claimed that the Dual Swin Transformer U-Net (DS-TransUNet) might be the first effort to simultaneously incorporate the advantages of hierarchical Swin Transformer into both encoder and decoder of the typical U shaped network to enrich the semantic segmentation quality. This model significantly outperforms especially in polyp segmentation task.

It is worth noting that color variation in histopathology images, cluster of nuclei with overlapping, shape, and size between different nuclei degrade the performance of the above mentioned nuclei-segmentation deep CNN models. In this study, we present a promising nuclei segmentation model for overcoming these limitations. To demonstrate the effectiveness of the model, we consider histopathology microscopic or fluorescent microscopic images of nuclei cells.

3 Methodology

Natural scenes have clear outlines that makes easy in segmentation, whereas medical images have great complication itself causes the separation between its nucleus to be blurred, not clear enough. To visualize the structure of tissues in stains, different color issues makes trouble such as nuclei are in blue/purple and cytoplasm in pink. Other challenges are the characteristics of the nuclei such as shape, size, gray value, and clusters of nuclei with overlapping.

After studying several convolutional models, we noticed some flaws even though these models achieved impressive results. For example, insufficient deployment of shallow layer structure information (i.e. ENet), inadequate integration of high and low level semantic information (e.g. SegNet), limited size of receptive field (i.e. UNet). To overcome the above difficulties and challenges, we have proposed a new nuclei segmentation encoder-decoder model, namely, AlexSegNet where AlexNet used as an encoder base model.

3.1 Problem definition

Given a biological image set S, our motive is to assign each pixel of an image to a class belonging to either nuclei regions or non-nuclei regions. We have $S = \{I_j, G_j\}_{j=1}^{N}$, where I_j represents original RGB images, and $G_j = \left\{ g_i^{(j)}, i = 1, 2, \ldots, |I_j|, g_i^{(j)} \in \{0, 1\} \right\}$ denotes their corresponding ground truth mask results. Each mask has each pixel labelled as $g_i^{(j)} = 1$ for representing nuclei regions and $g_i^{(j)} = 0$ for representing non-nuclei regions. The aim is to train a pixel-wise classifier to learn the following mapping function:

$$S_j = AlexSegNet(I_j) \tag{1}$$

where *AlexSegNet* represents our proposed model and S_j is the segmented output.

3.2 Proposed architecture

The overall proposed architecture consists of three levels, as shown in Fig. 2. We will continue next sub-sections with discussions of levels.

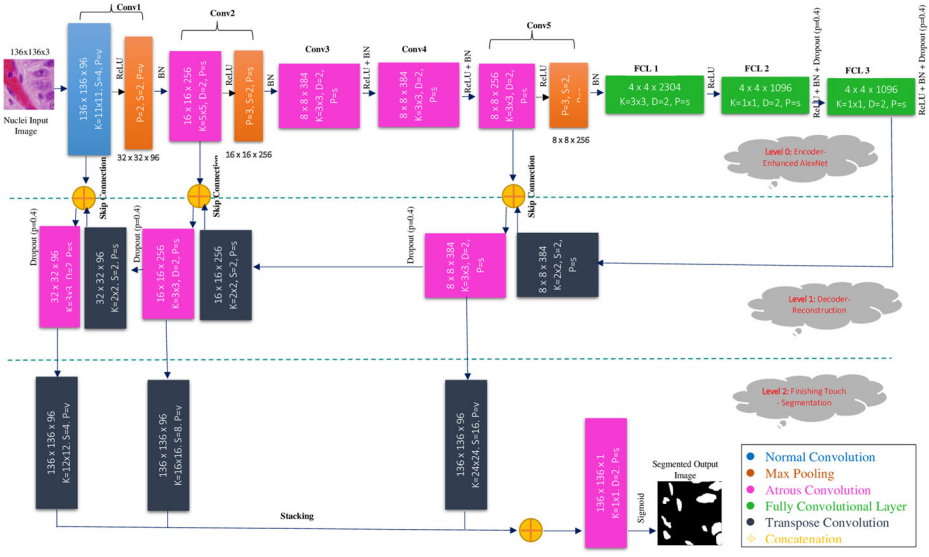


Fig. 2 Our Proposed Architecture, namely, AlexSegNet. K: kernel, S: Stride, D: Dilated rate, P: Padding, s: same, v: valid

3.2.1 Level 0: Encoder – Enhanced AlexNet

AlexNet, one of the first CNN network used for classification purpose. Till date no one explored this engrained network for medical image segmentation. The network have multi-size kernels (11 × 11 at conv1, 5 × 5 at conv2, 3 × 3 at rest of network layes) makes larger receptive field which increases the perception ability of cluster nucleus and will handle under-segmentation. The network consists of eight layers: five convolutional layers and three fully-connected layers. We have changed few thing in internal structure of network to enhance the performance.

In the convolutional blocks, the replacement of normal convolution (blue block) with atrous convolution (pink block) has done. The reason behind replacement is exponential expansion of reception field which is a promising side to handle overlapped cluster nucleus segmentation.

Atrous convolution is support exponential expansion of the receptive field without loss of resolution. It is applied to input feature map ($F : Z^2 \rightarrow R$ be a feature map discrete function) with definite gaps in the kernels ($k : \Omega_r \rightarrow R$ be a kernel of size $(2r + 1)^2$). Atrous convolution can be formulated as [42].

$$(F *_a k) = \sum_i \sum_{aj} F(i)k(aj) \tag{2}$$

where 'a' be a atrous factor. If atrous rate is 1, it means the convolution kernel is normal, and if the atrous rate is 2, then there is a skip of one pixel per input. Increasing the stride reduces the dimension of the output. A 2 × 2 atrous convolution has the same receptive field as a 3 × 3 un-atrous convolution. It will be better to visualize how a filter with spread out receptive field as in Fig. 3:

AlexNet is generally used for classification purpose where convolution blocks are usually flattened and then fully connected layers end up to give classification results. However, the retained spatial information got lost. Therefore, the three fully connected layers has been converted to fully convolutional layers (green block) as FCL 1, FCL 2, and FCL 3, as shown in Fig. 2. These layers are very essential to retain micro information's which are very important especially for medical microscopic imaging. If a fully connected (FC) vector (f_c^p) consists of p number of neurons, then the corresponding fully convolutional layer (f_{cl}) can be mapped as follows:

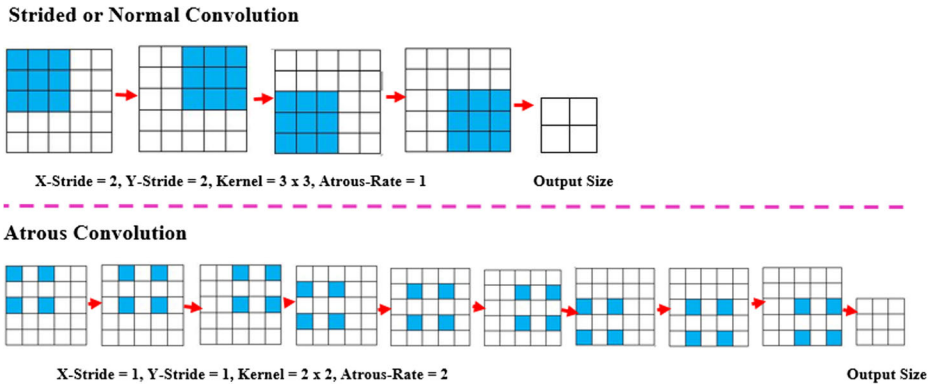


Fig. 3 Reasonable description between normal and atrous convolution

$$f_c^p \rightarrow^{l \times l,\quad d=2} [f_{cl}]_{i \times j \times p} \tag{3}$$

Where $l \times l$ is the kernel size and d is the dilation rate. This mapping of fully connected to fully convolutional layer creates a tensor of size $i \times j$ with p number of channels.

FCL 1 is convoluted over 2304 channels of kernel size 3 × 3 with atrous rate 2. Similarly, FCL 2 and 3 convoluted over 1096 channels of kernel size 1 × 1 with atrous rate 2. These FCLs gone through dropout with probability 0.4 to reduce overfitting issues.

3.2.2 Level 1: Decoder – Reconstruction

In encoder network, we reduces size of feature maps gradually up to 4 × 4 to extract low level textural information. Now, the decoder network will be gradually reconstruct the feature maps by up-sampling high level information. Every step in the decoding module consists of three details. First, an up-sampling transpose convolution operation enlarges the size of every feature map twice. Second, the feature maps concatenate via skip connection with the corresponding feature maps from the encoding network to avoid losing pattern or spatial information. The concatenation can fuse the low and high level information of the feature maps, and enhance the perception ability to smaller and larger nucleus. Third, it conducts one 3 × 3 atrous convolution operation with rate 2 followed by a ReLU and dropout operations.

In skip connection operation, the skip connection explicitly concatenate the feature maps generated in upper level (ul) with current level (cl) feature maps. Let $\sum_{cl}^{ul-i}$ be concatenate layer. The convolutional (normal or atrous) feature map $\mathbb{C}^{cl-1}$ in the preceding layer is up-sampled by a scale factor φ_s (where $\varphi_s = 2^n$) which will increase the dimension of the $(cl-1)^{th}$ layer by a factor of 2^n, and then concatenate it with an upper level convolutional feature layer C^{ul-i} where i be the number of skipped layers from the concatenated layer. It can be formulated as

$$\sum_{cl}^{ul-i} = \left(\mathbb{C}^{cl-1} * \varphi_s\right) \oplus \left(C^{ul-i}\right) \tag{4}$$

There are three up-sampling feature maps. First up-sampled feature map of size 8 × 8 with channels 384 has concatenate with conv5 block. Second up-sampled feature map of size 16 × 16 with channels 256 has concatenate with conv2 block. Third up-sampled feature map of size 32 × 32 with channels 96 has concatenate with conv1 block.

3.2.3 Level 2: Final touch – Segmentation

As a last contribution, we have also introduced a stacked network where feature maps are stacks on top of each other. By inspired from PSPNet [51], the third level is presented where three different scales of convoluted layers from level 1 are up-sampled again as original image resolution 136 × 136 with 96 channels each. The stacking via concatenation of these newly un-sampled feature maps comprises of global and local information which will be able to capture the context of the whole image. For final representation, a last 1 × 1 atrous convolution with 1 channel operation has applied to produce the per pixel prediction i.e. the segmented mask S_j.

As our problem definition is of two class labels i.e. binary class. So we have considered a sigmoid function as it reflects real values ranging between 0 and 1. The sigmoid activated output function P_r may be denoted as

$$P_r\left(O_j = q|S_j;\rho\right) = \frac{1}{1+\sum_{q=0}^{1}\exp\left(-\rho_q S_j\right)} \tag{5}$$

where O_j be the actual output belonging to q number of classes ranging between 0 and 1, S_j is the final output feature map obtained at last atrous convolutional layer and ρ be the parameter for the sigmoid function.

4 Experimental analysis

This section is divided into three subsections. In subsection (4.1), we analyze the performance of the proposed model, namely, AlexSegNet over several optimization strategies on the 2018 Data Science Bowl [46] and TNBC [49] datasets via performance metrics. In subsection (4.2), we compare proposed network model against existing base state-of-the-art segmentation CNN models. In subsection (4.3), assessment of the proposed model has been studied along with recently reported article results.

4.1 Brief introduction to datasets

For experimental evaluation, the 2018 Data Science Bowl [35] has been chosen which consists of 37,333 manually annotated nuclei in 841 2D images from different samples. The nuclei are derived from several organisms including humans, mice, and flies. As well, nuclei have been captured in a variety of conditions such as stains of fluorescent and histology, color of tissues, several magnifications, and illumination effect variations. The training dataset consists of total 670 images along with ground-truth. We have not considered test dataset (total of 68 images) since it doesn't contains of ground-truth. From training dataset, we splits 600 for training and rest 70 for testing. The dataset have imbalance distributions of samples of categorical nuclei, as shown in Fig. 4.

The second dataset has been selected is Triple Negative Breast Cancer (TNBC) nuclei segmentation database [11]. It contained 50 H&E stained samples those taken at 40x magnification with 512 × 512 resolution images. All 50 images were extracted from 11 number of patients with several cell types including normal epithelial cells, myoepithelial breast cells, fibroblasts, invasive carcinoma cells, endothelial cells, macrophages, adipocytes, and inflammatory cells. Overall summary for both datasets are listed in Table 2.

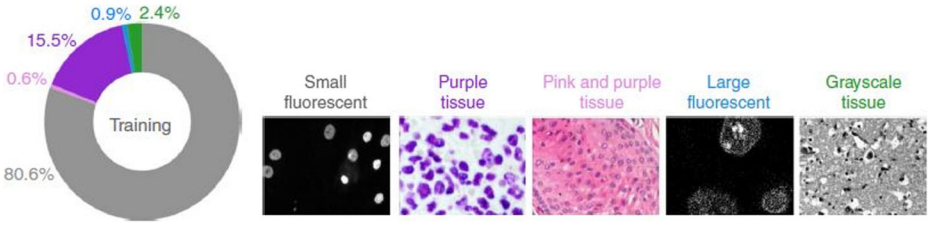


Fig. 4 Dataset sample distributions and variety of sample conditions [37]

Table 2 Summary for datasets Data Science Bowl and TNBC

Dataset	No. of Images	Resolution	Training Size	Testing Size	No. of Images after Augmentation
Data Science Bowl	670	256×256 to 1024×1024	600	70	600×5=3000
TBNC	50	512×512	47	3	47×5=235

4.1.1 Data augmentation

Since the used training dataset has a low number of samples, the proposed model might be prone to over-fit problem. To produce a proficient performance by addressing over-fitting challenge, data augmentation procedures are essential. Data augmentation through sample rotation by 50 degree, shear with factor 0.5, zoom with factor 0.2, fill mode with reflect, and lastly shift by factor 0.2 both width and height. That allow to extending the size of the dataset without deteriorating its quality. After these five approaches of data augmentation, the total number images listed in Table 2.

4.1.2 Evaluation metric

Specificity (Spec), Precision (Pre), Recall (Rec), F_1-score (F_1), Matthews Correlation Coefficient MCC, and Accuracy (Acc) are employed as the assessment metric for nuclei segmentation. The calculation formulas are shown in Eqs. (6–10) [40]. Spec represents the truely negative segmented nuclei rate, Acc represents the ratio of the truely positive segmented nuclei and all the background pixels. Among the total amount of segmented nuclei pixels, Precision represents the ratio of truley positive predictive segmented nuclei in label images, and Rec represents the percentage of the total amount of nuclei pixels correctly segmented in label images. F_1-score value is used to estimate the harmonic average of the Prec and Rec. F_1 could give a biased outcome since it doesn't include TN, and in such cases MCC is a perfect balance metric for evaluating performance. TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative respectively.

$$Spec = \frac{TN}{TN + FP} \tag{6}$$

$$Rec = \frac{TP}{TP + FN} \tag{7}$$

$$Pre = \frac{TP}{TP + FP} \tag{8}$$

$$F_1 = \frac{2 \times Pre \times Rec}{Pre + Rec} \tag{9}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{10}$$

4.1.3 Training and testing parameters

We trained the proposed model for 20 epochs and 14 epochs on the training dataset from 2018 Data Science Bowl and TNBC nuclei images. From this training samples, the splits for training-validation has done at 0.9:0.1 ratio. The overfitting avoiding strategy has done through data augmentation. Throughout training, a batch size of 10 used, and the loss estimating is binary cross entropy. It can be seen from Fig. 5 that when the iteration period is about 20 for Data Science Bowl and about 14 for TNBC, the validation loss/accuracy stops decreasing/increasing which means of leading to the training termination. In fact, the model starts to get relatively decent performance when training after the few epochs in case of Data Science Bowl dataset. However, this decent performance is not seen in case of TNBC due to very limited number of histopathology nuclei sample images.

4.2 Evaluation of proposed model on 2018 data science bowl dataset and TNBC dataset

Table 3 shows the performance results of the nuclei segmentation of proposed model on 2018 Data Science Bowl dataset. These results are measured in terms of several well-known metrics

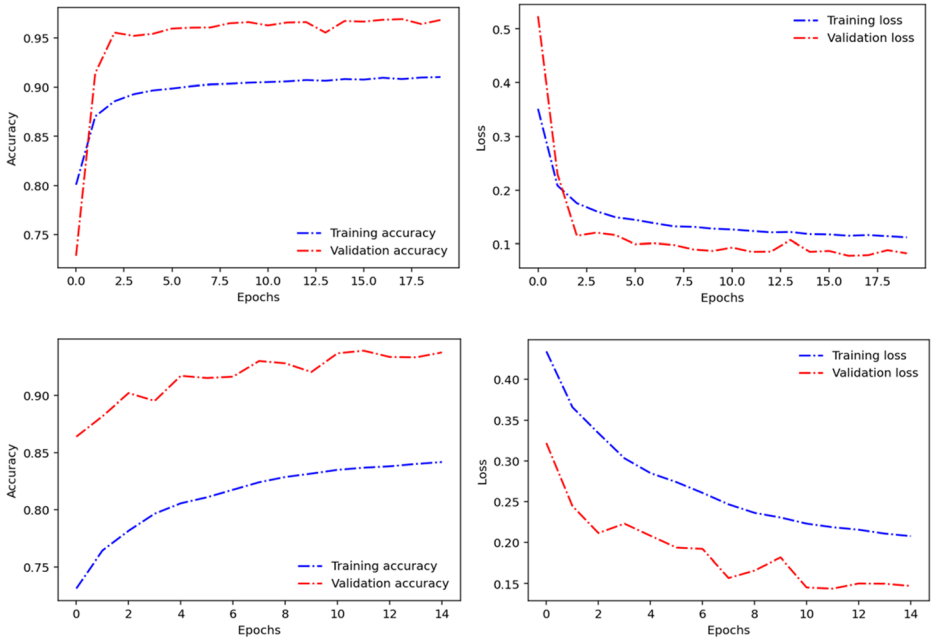


Fig. 5 Loss and accuracy curve of training and validation on 2018 Data Science Bowl dataset (first row) and TNBC Nuclei Segmentation dataset (second row) under AlexSegNet proposed model

Table 3 Performance Metric values for proposed model, AlexSegNet, on 2018 Data Science Bowl Dataset

		adamax	adam	nadam	rmsprop	sgd	adagrad	adadelta
Small Fluorescent	Spec	0.9857	0.9854	0.9775	0.9842	0.9410	0.9119	0.9129
	Pre	0.9217	0.9031	0.8797	0.9068	0.6873	0.5129	0.5232
	Rec	0.9132	0.9275	0.9096	0.9313	0.6917	0.5421	0.5456
	F_1	0.9150	0.9123	0.8911	0.9166	0.6834	0.5270	0.5341
	MCC	0.8990	0.8963	0.8704	0.9007	0.6252	0.5012	0.5134
	Acc	0.9738	0.9732	0.9659	0.9740	0.9074	0.8912	0.8876
Purple Tissue	Spec	0.9546	0.9409	0.9106	0.9495	0.8412	0.8076	0.8012
	Pre	0.9188	0.9017	0.8583	0.9114	0.7477	0.5487	0.5223
	Rec	0.7848	0.8329	0.7352	0.7870	0.6247	0.4231	0.4145
	F_1	0.8427	0.8617	0.7686	0.8374	0.6427	0.4777	0.4621
	MCC	0.7746	0.7949	0.6820	0.7699	0.5000	0.4043	0.3976
	Acc	0.9033	0.9111	0.8680	0.9027	0.7921	0.6987	0.6791
Large Fluorescent	Spec	0.9748	0.9546	0.9139	0.9749	0.8922	0.8434	0.8509
	Pre	0.7611	0.6917	0.6159	0.7619	0.5661	0.4986	0.4899
	Rec	0.8779	0.9228	0.9300	0.9129	0.9000	0.6025	0.6265
	F_1	0.8152	0.7892	0.7266	0.8303	0.6732	0.5456	0.5498
	MCC	0.7987	0.7747	0.7144	0.8176	0.6612	0.5257	0.5178
	Acc	0.9664	0.9537	0.9204	0.9699	0.9008	0.8201	0.8147
Pink Purple Tissue	Spec	0.9860	0.9733	0.8814	0.9870	0.9350	0.8753	0.8852
	Pre	0.9227	0.8591	0.6266	0.9239	0.7197	0.5980	0.5982
	Rec	0.6876	0.6705	0.8203	0.6486	0.6885	0.4321	0.4147
	F_1	0.7880	0.7531	0.7105	0.7621	0.7038	0.5016	0.4898
	MCC	0.7573	0.7101	0.6377	0.7326	0.6341	0.4769	0.4732
	Acc	0.9278	0.9142	0.8695	0.9210	0.8868	0.6726	0.6509
Grayscale Tissue	Spec	0.9957	0.9506	0.9459	0.9760	0.9915	0.8999	0.8839
	Pre	0.6456	0.5076	0.4158	0.5661	0.5432	0.3515	0.3523
	Rec	0.5521	0.6018	0.4550	0.3705	0.4331	0.2924	0.2848
	F_1	0.5951	0.5507	0.4345	0.4479	0.4819	0.3192	0.3149
	MCC	0.5715	0.5113	0.3848	0.4221	0.4702	0.3012	0.2934
	Acc	0.9252	0.9233	0.9075	0.9287	0.9215	0.8843	0.8765

such as specificity, precision, recall, F_1 score, MCC, and accuracy. The tested nuclei samples are categories into small fluorescent, purple tissue, large fluorescent, pink purple tissue, and grayscale tissue. The proposed model evaluation also analysed via several optimization strategies such as adamax, adam, nadam, rmsprop, sgd, adagrad, and adadelta. Now, we will investigate each nuclei sample categories.

In small fluorescent nuclei samples, the metric values of specificity (0.9857, 0.9854, 0.9842), F_1 (0.9150, 0.9123, 0.9842), and MCC (0.8990, 0.8963, 0.9007) gets approximately equal in adamax, adam, and rmsprop optimizers, respectively. We have received a little lower in nadam optimizer. We also noticed that the recall and precision has balanced values in adamax than adam, rmsprop, and nadam. Here, the other optimizers like sgd, adagrad, and adadelta are not giving promising results, and adamax is most promising.

In case of purple tissue samples, the specificity (adamax = 0.9546, adam = 0.9409, rmsprop = 0.9495), F_1 (adamax = 0.8427, adam = 0.8617, rmsprop = 0.8374) and MCC (adamax = 0.7746, adam = 0.7949, rmsprop = 0.7699) gets nearly equal but slightly lesser than small fluorescent samples. The nadam optimizer also promising. The recall and precision values difference is somehow bigger as compared to small fluorescent samples which as consequences shows lower F_1 scores. The recall values are low due to improper segmentation of region interest (ROI) from nuclei. The sgd, adagrad, and adadelta as usual shows low performances.

In large fluorescent nuclei samples, we received high specificity values over adamax, adam, rmsprop as usual. The proposed model is capable enough to truly discriminate the negative portions from the ROI nuclei portion. The recall and precision unbalance factor also observed like the case of purple tissue but opposite i.e. precision values are low and recall values are high. That means, the proposed model segmenting ROI nuclei very well which produces good recall value, along with some unwanted segmented results which reduces the precision values. Among promising optimizers, here we noticed rmsprop is slightly better.

In case of pink purple tissue, we have faced similar unbalance effects in recall and precision values where recall values are low. This is happening due to different color of tissues in histopathology stains that makes improper segmentation of tissue. In this case, the sgd optimizer shows a balanced factor between recall and precision however low metric results. In terms of F_1 and MCC, the adamax is most promising.

In grayscale tissue, we have received maximum specificity of 99.57% through adamx optimizer. However, we also received lowest performances over all optimizer and metrics.

As a whole, the adamax shows outperform over all optimizer. Therefore, all succeeding evaluations are continued with adamax. In terms of accuracy, the small fluorescent samples shows highest accuracy of 97%, followed by large fluorescent of 96%.

Now, Fig. 6 shows the performance results of the nuclei segmentation of proposed model on TNBC dataset. These results are also measured in terms of several metrics and optimization strategies. The tested nuclei samples are only category of breast histopathology tissues. In these nuclei samples, the metric values of specificity gets highest in adamax, nadam, adagrad and adadelta optimizations. This is only the case where adagrad and adadelta achieved impressing results. We are received good precision values in this dataset where adamax, nadam and rmsprop shows promising. The precision values are good means our model doing well segment without much wrongly positive nuclei cells. The recall is not as per expectation due to lesser number of image samples in this dataset. The adamax optimization giving highest recall value. Due to the less recall values, the corresponding F1 score and MCC also gets effected. The achieved accuracy of our proposed model is approximately 90% through adamax and rmsprop optimization strategies.

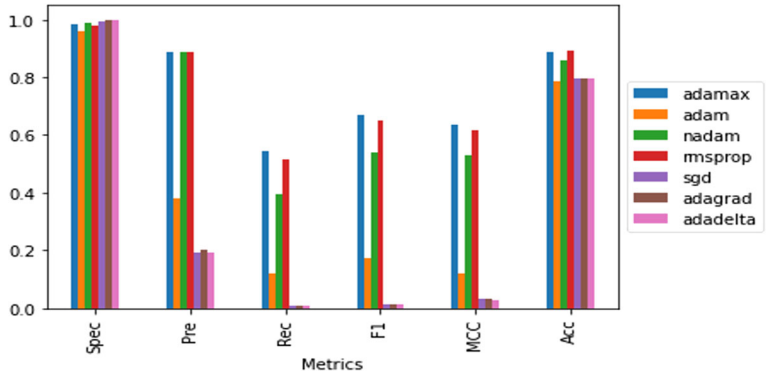
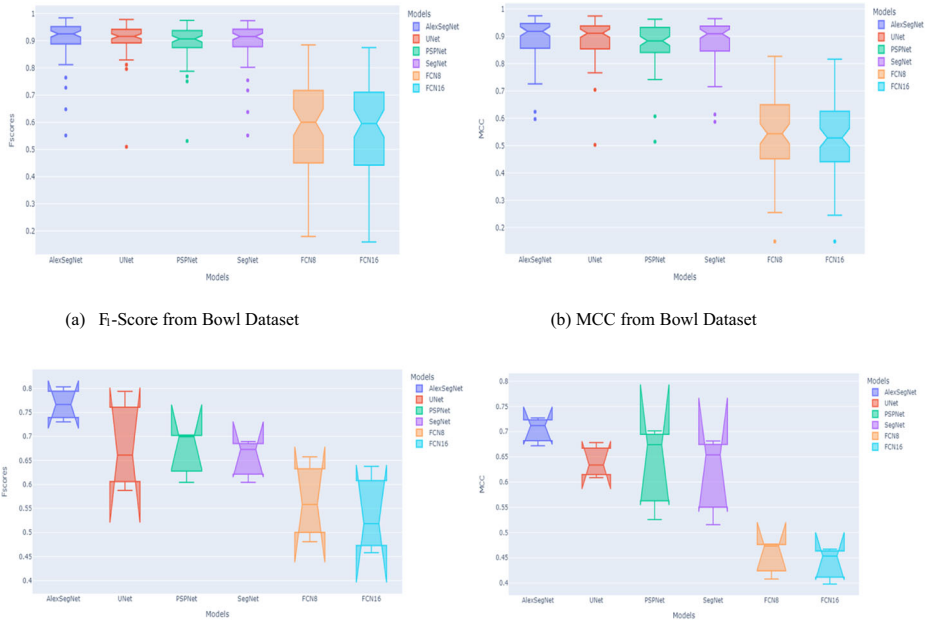


Fig. 6 Performance Metric values for proposed model, AlexSegNet, on TNBC Nuclei Segmentation Dataset

4.3 Comparative study of proposed model AlexSegNet with base state-of-the-art models

In this section, we compare AlexSegNet against UNet [44], PSPNet [51], SegNet [4], and FCN [28] versions 8 and 16. All models have their own design of architectures whereas FCNs are used base model VGG for encoder. Therefore, we have used AlexNet as base model for our proposed segmentation model.

Figure 7 shows the boxplots of F1-measure and MCC for all nuclei segmentation models. A boxplot is analysed through a given scores of test images with a specific prototypical five-number of definition summary such as the maximum (max), the minimum (min), the sample median (middle horizontal line), the first quartiles (q1), and third quartiles (q3). As in the Fig. 7a and b shows for Data Science Bowl dataset, the FCN 8 and FCN 16 models have not any outliers on F1-score and one outliers on MCC values, however, both the models shows lowest performance values of sample median, min, max, q1 and q3. The same scenarios has also noticed in Fig. 7c and d for TNBC dataset. In case of Data Science Bowl dataset, U-Net and SegNet models has the approximately equal median F1-score and MCC, but F1-scores produces the higher number of outliers than MCC evaluation metric values. However, in case of TNBC dataset, SegNet shows higher mean values than U-Net for both F1-score and MCC. Whereas PSPNet model promising as a fourth best model for Data Science Bowl dataset and second best model for TNBC dataset in our comparative evaluation. As we have seen for Data Science Bowl dataset, AlexSegNet model almost achieved the highest median F1-score of 95% with 4 outliers and MCC of 94% with 2 outliers. In TNBC dataset, AlexSegNet model



(a) F_1-Score from Bowl Dataset

(b) MCC from Bowl Dataset

(c) F_1-Score from TNBC Dataset

(d) MCC from TNBC Dataset

Fig. 7 Boxplots of F1-score and MCC of the six nuclei segmentation models where (**a**) and (**b**) for 2018 Data Science Bowl Dataset, (**c**) and (**d**) for TNBC Nuclei Segmentation Dataset

achieved the highest median F1-score of 77% and MCC of 71%. The TNBC based boxplots are looks unusual due to its lesser number of testing samples.

To provide a better visual understanding of the segmentation results, typical segmented outputs of nuclei from different categorical samples are shown in Fig. 8 under various base state-of-the-art models. In case of grayscale tissue samples, the PSPNet shows most favourable segmented output.

4.4 Comparative assessment of proposed model AlexSegNet with recently reported state-of-the-art article models

Table 4 shows quantitative comparative assessment of the proposed model along with recently reported article results. We compared our model against several deep learning based medical image segmentation models listed in Table 4. Our model performs competitively with present state-of-the-art methods on the two datasets i.e. 2018 Data Science Bowl and TNBC in the integrity of the segmentation of a multiple nucleus. In order to make the comparison objectively, we followed the articles proposed models for two datasets separately. In case of 2018.

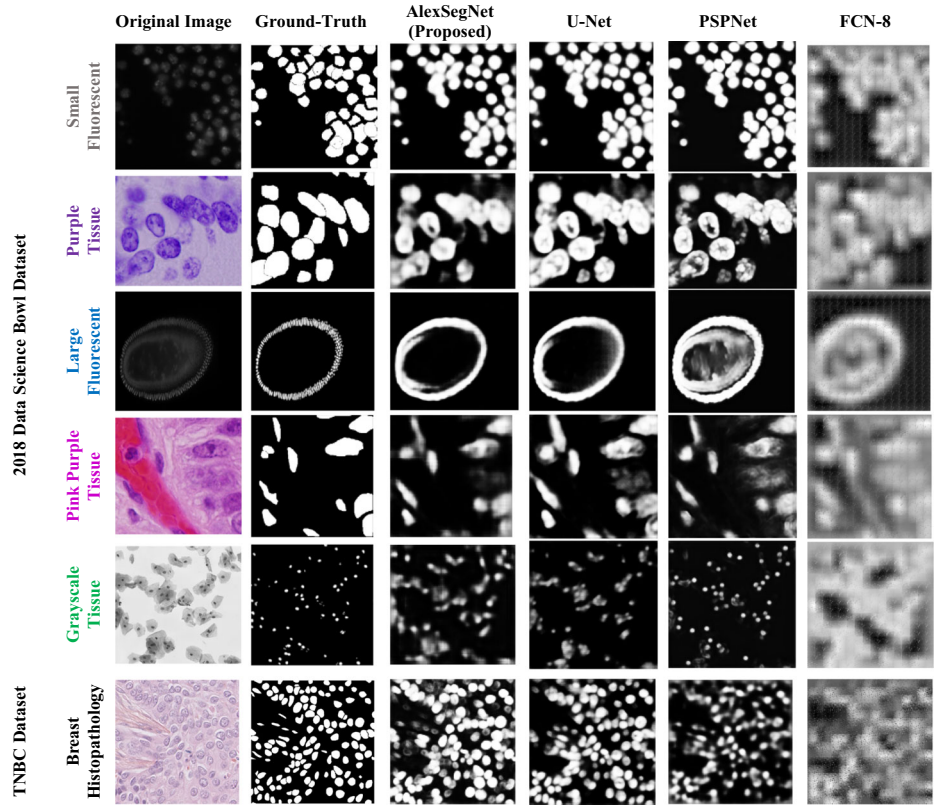


Fig. 8 Typical segmentation results of various nuclei samples from 2018 Data Science Bowl dataset and TNBC dataset. Column (1) shows original samples, column (2) shows ground-truth binary mask, column (3) shows segmented results under proposed model, column (4) shows segmented results under U-Net model, column (5) shows segmented results under PSPNet model, and last column shows segmented results under FCN-8 model

Table 4 Quantitative comparison of different methods applied to the 2018 Data Science Bowl and TNBC datasets. + indicates chosen best metric values

Dataset	Article Models	Recall	Precision	F1-Score
2018 Data Science Bowl	DoubleU-Net [1], 2020	0.6407	0.9496	–
	FANet [43], 2022	0.9222	0.9194	0.9176
	Attention U-Net [38], 2018	0.9183	0.9235	0.9179
	MSRF-Net [2], 2022	0.9402	0.9022	–
	DS-TransUNet-L [50], 2022	0.9378	0.9124	0.9219
	DCSAU-NET [9], 2022	0.9221	0.9063	0.9083
	MKDCNet [15], 2022	0.9270	0.9194	0.9237
	DeepLabV3+ (ResNet50) [39], 2018	0.9220	0.8902	0.9134
	PraNet [34], 2020	0.9182	0.8438	–
	Ours+	0.9313	0.9239	0.9166
TNBC	DeconvNet [22], 2015	0.773	0.864	0.805
	Ensemble [18], 2017	0.900	0.741	0.802
	Two-stage learning U-Net (DLA) [19], 2019	0.833	0.826	0.829
	SU-Net [12], 2020	0.853	0.792	0.806
	GB U-Net [23], 2021	0.5541	0.8102	0.6581
	Ours+	0.5421	0.8863	0.6688

Data Science Bowl dataset, the network models listed are DoubleU-Net [1], FANet [43], Attention U-Net [38], MSRF-Net [2], DS-TransUNet-L [50], DCSAU-NET [9], MKDCNet [15], DeepLabV3+ (ResNet50) [39], and PraNet [34]. The performance results of the networks utilized in this comparison are sourced from the respective publication. In terms of recall values, our model achieved very well metric value of 0.9313 which is just few percentage less than newly proposed models in 2022 i.e. MSRF-Net [2] and DS-TransUNet-L [50]. In precision values, our model shows second best achievement with 0.9239 where DoubleU-Net shows 0.9496. The results of the comparison confirmed the competitive of our model which achieved a F_1-score of 0.9166 for 2018 Data Science Bowl dataset. This outcome demonstrates that our model has a great generalization ability since these segmented nuclei images having accurately segmented with lesser wrongly negative nuclei (recall) and wrongly positive nuclei (precision) cells.

In case of TNBC dataset, the network models listed are DeconvNet [22], Ensemble [18], two-stage learning U-Net (DLA) [19], two-stage stacked U-Net (SU-Net) [12], and gradient boosting network U-Net (GB U-Net) [23]. The results of the comparison confirmed the superiority of our model which achieved a precision of 0.8863. Our model highly effected in recall value which consequences in F_1-score as well even after highest precision value. This outcome also demonstrates that our model has a generalization ability since these segmented nuclei images having average wrongly negative nuclei (recall) and lesser wrongly positive nuclei (precision) cells.

5 Computational cost or time complexity analysis

The proposed AlexSegNet model is primarily based on three levels i.e. level-0, level-1, and level-2 in parallel. In this subsection, we will compute the computational cost from each level. Our proposed model framework has convolutional layers (either normal or atrous), fully convolutional layers, pooling layers, transpose convolutional layers, concatenations, batch normalizations, and dropouts. However, the pooling layers, concatenations, dropouts, and

batch normalizations only consume 5% to 10% of the computational time [53], while convolutional layers, fully convolutional layers, transpose convolutional layers consume the majority of the time. So to simplify, we will only elaborate the complexity of these layers as in [29].

Let p be a convolutional layer, so the complexity at this layer would be $Big{-}O\left(\frac{1}{\alpha}\left[I_{p-1}\times K_p\times S_p^2\times F_p^2\right]\right)$ where $I_{p\ -\ 1}$ is the number of input channels in the $(p\ -\ 1)^{th}$ layer, K_p is the number of kernels in the p^{th} layer, S_p^2 is the spatial size of each kernel in the p^{th} layer, F_p^2 is the spatial size of output map, $\frac{1}{\alpha}$ is the feature map size reduction rate after pooling execution at each convolutional layer. Therefore, the summing up the time complexity for all the $\aleph = 9$ (level-0: 5 convolutional layers either normal or atrous, level-1: 3 atrous convolutional layers, level-2: 1 atrous convolutional layer) number of convolutional layers in the framework, we get

$$\copyright = Big{-}O\left(\sum_{p=1}^{\aleph}\frac{1}{\alpha}\left[I_{p-1}\times K_p\times S_p^2\times F_p^2\right]\right) \tag{11}$$

Similarly, we will estimate the complexity for transpose convolutional ®or de-convolutional layers. Consequently, the calculative formula would be

$$® = Big{-}O\left(\sum_{p=1}^{\aleph}\beta\left[I_{p-1}\times K_p\times S_p^2\times F_p^2\right]\right) \tag{12}$$

where β is the feature map expansion rate after unpooling execution at transpose convolution layer. In case of Eq. (12), $\aleph = 6$ (level-1: 3 transpose convolutional layers, level-2: 3 transpose convolutional layers) number of transpose convolutional layers in the framework. Therefore, the overall complexity of our model is computed as

$$O = \copyright + ® \tag{13}$$

As consequences, our model uses 47 million parameter which is as compare lesser if we are comparing with FCN network [28]. However, the number of parameter used by UNet [44] is 2.6 million which consider as a limitation of the AlexSegNet that leads to an increase in the training time. In future, the number of parameters will be tried to reduce by replacing base AlexNet architecture kernel sizes. If we see from model size the point of view, AlexSegNet uses only 15 layers which is an advantage. Whereas, the U-Net [44] model consumed 23 layers and SegNet [4] utilized 26 layers where fully connected layers had been discarded.

6 Conclusion

This study introduces a deep convolutional encoder-decoder architecture based upon AlexNet, namely, 'AlexSegNet' for nuclei segmentation in microscopic pathological images. The architecture consists of three levels: Level 0 – enhanced AlexNet encoder, Level 1 – decoder for image reconstruction, and Level 2 – finishing touch for segmentation. The 2018 Data Science Bowl and TNBC datasets are used for experimentation and to demonstrate the proposed nuclei segmentation architecture. In the process of model training, numerous optimization have been analysed whereas adamax specified most promising results and improve the segmentation performance. From the comparative studies, it has been observed that only

grayscale tissue samples segmentation has not been done well by our model whereas PSPNet state-of-the-art model well performed in these grayscale samples. As overall evaluation reveals that our proposed architecture achieves maximum recall value of 0.9313, precision value of 0.9239, F1 score of 0.9166 for Data Science Bowl dataset, and in case of TNBC dataset, recall value of 0.5421, precision value of 0.8863, and F1 score of 0.6688.

In future, the number of parameters will be tried to reduce by replacing base AlexNet architecture kernel sizes. The proposed AlexSegNet architecture will also be enhanced in many image processing handcraft features point of view such as Fourier transform, wavelet transform, and so on.

Acknowledgements The work presented here is being conducted in the Department of CSE, SRM IST, Delhi-NCR Campus, Ghaziabad, India. The authors would also like to thank HOD, Dept. of CSE, SRM IST for his kind support to carry out this work.

Data availability Data sharing not applicable to this article as no datasets were generated during the current study.

Declarations

Conflict of interests The author declares no potential conflict of interest with respect to the authorship and/or publication of this article.

References

1. Abhishek S, Debesh J, Chanda S, Pal U, Johansen HD, Johansen D, Riegler MA, Ali S, Halvorsen P (2022) MSRF-Net: a multi-scale residual fusion network for biomedical image segmentation. IEEE J Biomed Health Inform. https://doi.org/10.48550/arXiv.2105.07451
2. Ailiang L, Bingzhi C, Jiayu X, Zheng Z, Guangming L, David Z (2022) "DS-TransUNet: dual Swin transformer U-net for medical image segmentation. IEEE Trans Instrum Meas 71. https://doi.org/10.1109/TIM.2022.3178991
3. Ambati LS, El-Gayar O, Nawar N (2021) Design principles for multiple sclerosis Mobile self-management applications: a patient-centric perspective. In: Proceedings of the 27th annual Americas Conference on Information Systems (AMCIS 2021)
4. Badrinarayanan V, Handa A, Cipolla R (2015) SegNet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. arXiv:1505.07293. [Online]. Available: https://arxiv.org/abs/1505.07293. Accessed 27 May 2015
5. Bakkouri I, Afdel K (2019) Computer-aided diagnosis (CAD) system based on multi-layer feature fusion network for skin lesion recognition in dermoscopy images. Multimed Tools Appl 79: 20483–20518
6. Bakkouri I, Afdel K, Benois-Pineau J, Catheline G (2022) BG-3DM2F: bidirectional gated 3D multi-scale feature fusion for Alzheimer's disease diagnosis. Multimed Tools Appl 81:10743–10776
7. Basavanhally A, Feldman M, Shih N, Mies C, Tomaszewski J, Ganesan S, Madabhushi A (2011) Multi-field of view strategy for image-based outcome prediction of multiparametricestrogen receptor-positive breast cancer histopathology: Comparison to Oncotype DX. J Pathol Inform 2. https://doi.org/10.4103/2153-3539.92027
8. Caicedo JC, Goodman A, Karhohs KW, Cimini BA, Ackerman J, Haghighi M, Heng CK, Becker T, Doan M, McQuin C, Rohban M, Singh S, Carpenter AE (2019) Nucleus segmentation across imaging experiments: the 2018 data science bowl. Nat Methods 16:1247–1253
9. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H (2018) Encoder decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV), pp 801–818

10. Cui Y, Zhang G, Liu Z, Xiong Z, Hu J (2019) A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images. Med Biol Eng Comput 57:2027–2043
11. PowersDMW (2010) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. Int J Mach Learn Technol 2:37–63
12. Debesh J, Michael AR, Dag J, Pal H, Havard DJ (2020) "DoubleU-Net: a deep convolutional neural network for medical image segmentation. In: Proceedings of IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS 2020)
13. Dubost F, Bortsova G, Adams H, Ikram A, Niessen WJ, Vernooij M, De Bruijne M (2017) GP-Unet: lesion detection from weak labels with a 3d regression network. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2017), pp 214–221
14. El-Gayar O, Ambati LS, Nawar N (2020) Wearables, artificial intelligence, and the future of healthcare. In: Strydom M, Buckley S (eds) In Book: AI and Big Data's Potential for Disruptive Innovation. IGI Global, pp 104–129
15. Fan D-P, Ji G-P, Zhou T, Chen J, Fu H, Shen J, Shao L (2020) PraNet: parallel reverse attention network for polyp segmentation. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2020), pp 263–273
16. Grahama S et al (2019) Hover-net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. Med Image Anal 58:1–14
17. Hassan L, Saleh A, Abdel-Nasser M, Omer OA, Puig D (2021) Promising deep semantic nuclei segmentation models for multi-institutional histopathology images of different organs. Int J Interact Multimed Artif Intell 6:35–45
18. He K, Sun J (2015) "Convolutional neural networks at constrained time cost. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR 2015), pp 5353–5360
19. Hipp JD, Fernandez A, Compton CC, Balis UJ (2011) Why a pathology image should not be considered as a radiology image. J Pathol Informat 2. https://doi.org/10.4103/2153-3539.82051
20. https://gco.iarc.fr/today/explore (2021) [Accessed on 10th December 2021]
21. 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
22. Kang Q, Lao Q, Fevens T (2019) "Nuclei segmentation in histopathological images using two-stage learning," In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A (eds) In Book: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019. Lecture Notes in Computer Science 11764:703–711
23. 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. Frontiers Bioeng Biotechnol 8. https://doi.org/10.3389/fbioe.2020.573866
24. Lagree A, Mohebpour M, Meti N, Saednia K, Lu FI, Slodkowska E, Gandhi S, Rakovitch E, Shenfield A, Sadeghi-Naini A, Tran WT (2021) A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks. Sci Rep 11: Article No. 8025
25. Lal S, Das D, Alabhya K, Kanfade A, Kumar A, Kini J (2021) NucleiSegNet: robust deep learning architecture for the nuclei segmentation of liver Cancer histopathology images. Comput Biol Med 128. https://doi.org/10.1016/j.compbiomed.2020.104075
26. Li X, Chen H, Qi X, Dou Q, Fu C-W, Heng P-A (2018) H-denseUnet: hybrid densely connected unet for liver and tumor segmentation from ct volumes. IEEE Trans Med Imaging 37:2663–2674
27. Li K, Yu L, Wang S et al (2020) Towards cross-modality medical image segmentation with online mutual knowledge distillation. Proc AAAI Conf Artif Intell 34:775–783
28. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of IEEE Conference Computer Vision and Pattern Recognition (CVPR 2015), pp 3431–3440
29. Lu L, Yang Y, Jiang Y, Ai H, Tu W (2018) Shallow convolutional neural networks for acoustic scene classification. Wuhan Univ J Natural Sci 23:178–184
30. Mehta S, Rastegari M, Caspi A, Shapiro L, Hajishirzi H (2018) ESPNet: Ef_cient spatial pyramid of dilated convolutions for semantic segmentation. arXiv:1803.06815. [Online]. Available: https://arxiv.org/abs/1803.06815. Accessed 25 Jul 2018
31. Mehta S, Mercan E, Bartlett J, Weave D, Elmore JG, Shapiro L (2018) Y-Net: Joint segmentation and classification for diagnosis of breast biopsy images. arXiv:1806.01313. [Online]. Available: https://arxiv.org/abs/1806.01313. Accessed 4 Jun 2018
32. Mehta S, Rastegari M, Shapiro L, Hajishirzi H (2019) ESPNetv2: A light-weight, power efficient, and general purpose convolutional neural network. arXiv:1811.11431. [Online]. Available: https://arxiv.org/abs/1811.11431. Accessed 30 Mar 2019

33. Naylor P, Laé M, Reyal F, Walter T (2017) "Nuclei segmentation in histopathology images using deep neural networks", Proceedings of IEEE Int. Symp. Biomed. Imag. (ISBI), pp. 933–936
34. Naylor P, Lae M, Reyal F, Walter T (2017) Nuclei segmentation in histopathology images using deep neural networks. In: Proceedings of 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp 933–936
35. Naylor P, Laé M, Reyal F, Walter T (2018) Segmentation of nuclei in histopathology images by deep regression of the distance map. IEEE Trans Med Imaging 38:448–459
36. Naylor P, Laé M, Reyal F, Walter T (2019) Segmentation of nuclei in histopathology images by deep regression of the distance map. IEEE Trans Med Imaging 38:448–455
37. Nikhil KT, Jha D, Riegler MA, Johansen HD, Johansen D, Rittscher J, Halvorsen P, Ali S (2022) FANet: a feedback attention network for improved biomedical image segmentation. http://arxiv.org/abs/2103.17235v3 [cs.CV]. Accessed 25 Mar 2022
38. Nikhil KT, Abhishek S, Ulas B, Jha D (2022) Automatic polyp segmentation with multiple kernel dilated convolution network. In: Proceedings of IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS), pp 317–322
39. Noh H, Hong S, Han B (2015) "Learning deconvolution network for semantic segmentation. In: Proceedings of the International Conference on Computer Vision (ICCV 2015), pp 1520–1528
40. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D (2018) Attention U-Net: learning where to look for the pancreas. In: Proceedings of 1st Conference on Medical Imaging with Deep Learning (MIDL 2018)
41. Paszke A, Chaurasia A, Kim S, Culurciello E (2016) ENet: A deep neural network architecture for real-time semantic segmentation. arXiv:1606.02147. [Online]. Available: https://arxiv.org/abs/1606.02147. Accessed 7 Jun 2016
42. Poplavskiy D (2018) Data Science Bowl—Discussion 55118 https://www.kaggle.com/c/data-science-bowl-2018/discussion/55118
43. Qing X, Zhicheng M, Na HE, Wenting D (2022) DCSAU-NET: a deeper and more compact split-attention U-NET for medical image segmentation. arXiv:2202.00972v2 [eess.IV] 24 Sep 2022
44. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. arXiv:1505.04597. [Online]. Available: https://arxiv.org/abs/1505.04597. Accessed 18 May 2015
45. Shah S (2020) Latest Statistics of Breast Cancer in India. Breast Cancer India. http://www.breastcancerindia.net/statistics/trends.html. Accessed 30 Aug
46. 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: Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pp 383–390
47. Vu QD et al (2019) Methods for segmentation and classification of digital microscopy tissue images. Front Bioeng Biotechnol 7:1–15
48. Xing F, Yang L (2016) Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. IEEE Rev Biomed Eng 9:234–263
49. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, Madabhushi A (2016) Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans Med Imag 35:119–130
50. Yu F, Koltum V (2016) Multi-scale context aggregation by Dilated Convolutions. arXiv:1511.07122v3 [cs.CV] 30 Apr 2016
51. Zhao H, Shi J, Qi X, Wang X, Jia J (2017) "Pyramid scene parsing network. In: Proceedings of IEEE Conference Computer Vision and Pattern Recognition (CVPR 2017), pp 2881–2890
52. Zhao H, Qi X, Shen X, Shi J, Jia J (2018) ICNet for real-time semantic segmentation on high-resolution images. arXiv:1704.08545. [Online]. Available: https://arxiv.org/abs/1704.08545. Accessed 20 Aug 2018
53. Zhou Y, Onder OF, Dou Q, Tsougenis E, Chen H, Heng P-A (2019) CIA-Net: Robust Nuclei Instance Segmentation with Contour-Aware Information Aggregation. In: Zhou Y, Chen H, Heng P-A, Onder OF, Tsougenis E, Dou Q (eds) In Book: Information Processing in Medical Imaging. IPMI 2019. Lecture Notes in Computer Science, vol 11492

Anu Singha is currently working as an Assistant Professor in the Department of Computer Science & Engineering, SRM Institute of Science and Technology (Formerly SRM University), Delhi-NCR Campus, ModiNagar. Prior to joining SRM University, he had worked for 1.5 years as Assistant Professor in Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research, Porur, Chennai. He was a Research Fellow (JRF+SRF) for 3 years on a DRDO funded project at Computer Vision Laboratory, Tripura University (A Central University). He had completed Ph.D. in April, 2021 after completion of M.Tech (CSE) degree from Tripura University (A Central University), 2015. He obtained another Masters' degree in Computer Applications (MCA) from South Asian University (A SAARC University), New Delhi in 2013. He completed bachelor degree as well in Computer Applications (BCA) from Maharaja Bir Bikram (MBB) College, Agartala in 2010. His topics of research interest are image processing, computer vision, medical imaging, deep learning, and machine learning.

Mrinal Kanti Bhowmik (Senior Member, IEEE) received the B.E. degree in computer science and engineering from the Tripura Engineering College in 2004, the M.Tech. degree in computer science and engineering from Tripura University (A Central University), India, in 2007, and the Ph.D. degree in engineering from Jadavpur University, Kolkata, India, in 2014. He has successfully completed two Department of Electronics and Information Technology (DeitY) (Now Ministry of Electronics and Information Technology (MeitY)) funded projects, one the Department of Biotechnology (DBT)-Twinning project, one Society for Applied Microwave Electronics Engineering and Research (SAMEER) funded project, and one Indian Council of Medical Research (ICMR) projectas the Principal Investigator. He is currently the Principal Investigator of a Defense Research and Development Organization (DRDO), Government of India funded project. Since July 2010, he has been serving with the Department of Computer Science and Engineering, Tripura University as an Assistant Professor. He was awarded the Short Term Indian Council of Medical Research (ICMR), Department of Health Research (DHR) International Fellowship from 2019 to 2020 as a Senior Indian Biomedical Scientist for bilateral cooperation in cross-disciplinary research area (i.e., biomedical diagnostic and inferencing systems). His current research interests are in the field of computer vision, security and surveillance, medical imaging, and biometrics.