


# Nucleus segmentation and classification using residual SE-UNet and feature concatenation approach incervical cytopathology cell images

Technology in Cancer Research & Treatment  
 Volume 22: 1-14  
 © The Author(s) 2023  
 Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
 DOI: 10.1177/15330338221134833  
[journals.sagepub.com/home/tct](https://journals.sagepub.com/home/tct)  


G Jignesh Chowdary, BTech<sup>1</sup>, Suganya G, PhD<sup>2</sup>, Premalatha M, PhD<sup>2</sup>, and Pratheepan Yogarajah, PhD<sup>3</sup> 

## Abstract

**Introduction:** Pap smear is considered to be the primary examination for the diagnosis of cervical cancer. But the analysis of pap smear slides is a time-consuming task and tedious as it requires manual intervention. The diagnostic efficiency depends on the medical expertise of the pathologist, and human error often hinders the diagnosis. Automated segmentation and classification of cervical nuclei will help diagnose cervical cancer in earlier stages. **Materials and Methods:** The proposed methodology includes three models: a Residual-Squeeze-and-Excitation-module based segmentation model, a fusion-based feature extraction model, and a Multi-layer Perceptron classification model. In the fusion-based feature extraction model, three sets of deep features are extracted from these segmented nuclei using the pre-trained and fine-tuned VGG19, VGG-F, and CaffeNet models, and two hand-crafted descriptors, Bag-of-Features and Linear-Binary-Patterns, are extracted for each image. For this work, Herlev, SIPaKMeD, and ISBI2014 datasets are used for evaluation. The Herlev dataset is used for evaluating both segmentation and classification models. Whereas the SIPaKMeD and ISBI2014 are used for evaluating the classification model, and the segmentation model respectively. **Results:** The segmentation network enhanced the precision and ZSI by 2.04%, and 2.00% on the Herlev dataset, and the precision and recall by 0.68%, and 2.59% on the ISBI2014 dataset. The classification approach enhanced the accuracy, recall, and specificity by 0.59%, 0.47%, and 1.15% on the Herlev dataset, and by 0.02%, 0.15%, and 0.22% on the SIPaKMeD dataset. **Conclusion:** The experiments demonstrate that the proposed work achieves promising performance on segmentation and classification in cervical cytopathology cell images..

## Keywords

Cervical cancer, pap smear, deep learning, segmentation, transfer learning

Received: May 4, 2022; Accepted: September 30, 2022.

## Introduction

Cervical cancer is one of the leading causes of death in women across the world<sup>1</sup>. The treatment and survival chances from this cancer are heavily dependent on the stage at the time of diagnosis. When cervical cancer is diagnosed at an early stage (or pre-cancerous stage)<sup>1</sup>, the chances of survival are very high, and the recovery is also fast. Cervical cytology is the most popular and trusted procedure for the screening of cervical cancer at early stages. This is a physical procedure in which few cells are collected from the cervix and are transferred into a container with special liquid (in case of liquid-based pap smear) to preserve the sample or onto a glass slide (in case of conventional pap smear) for examination under a microscope. This procedure has shown promising results in reducing the mortality rate of cervical cancer in women<sup>2</sup>. This procedure was performed for the screening of cervical cancer across the world. However, this

procedure is not available for population-wide screening in underdeveloped and developing countries because of its complexity and tedious nature as it involves a human intervention to manually examine for the abnormal cells in the cytology specimen<sup>3</sup>. But the automation of this examination procedure with computerized techniques like Artificial Intelligence (AI) will increase the efficiency and also reduce the detection time<sup>4</sup>.

Over the last few decades, there has been a lot of research that is done in automating several medical practices with AI via

<sup>1</sup> Stony Brook University, New York, US

<sup>2</sup> Vellore Institute of Technology, Chennai, India

<sup>3</sup> Ulster University, Northern Ireland, UK

### Corresponding Author:

Pratheepan Yogarajah, Ulster University, Northern Ireland, UK.

Email: [p.yogarajah@ulster.ac.uk](mailto:p.yogarajah@ulster.ac.uk)



machine learning and deep learning<sup>5</sup>. These methods have shown promising results in the diagnosis of Pneumonia, Brain tumors, Heart diseases, COVID-19, Tuberculosis<sup>6</sup>, and also in the diagnosis of other cancers like breast<sup>7</sup>, lung, and brain. Even in this field, there are several studies that were proposed for the screening of cervical cancer from cervical cytology images<sup>3,8-10</sup>. Like, Dong et al.<sup>11</sup> proposed an approach that uses a canny segmentation algorithm to segment the nuclei regions from the cytology images from the single-cell dataset; from these regions edge features, are extracted using adaptive gradient vector flow snake model, and these features are used to train the support vector machine algorithm for classifying normal and abnormal cells. Few other studies were presented in<sup>12,13</sup> using the same dataset, where the authors used Fuzzy C means clustering and Radiating-Gradient-Vector Flow (GVF) model for nuclei segmentation. Marinkis et al.<sup>14</sup> performed Benign/Malignant classification on the same dataset using nearest neighbor classifiers trained with features selected using genetic algorithms.

Genctav et al.<sup>15</sup> implemented smear level segmentation based on circularity, uniformity, and nuclear size. In the later part of the study, they also used an unsupervised learning approach to conduct binary classification on a smear level dataset. Their results show improved effectiveness when dealing with challenges associated to poor strained quality. Bora et al.<sup>16</sup> used shape-based nuclei features extracted by the Maximal stable external region (MSER) algorithm followed by thresholding ratio and some morphological operations for smear level segmentation. To analyse the hyperchromatic variations in the nuclei, the scientists employed textual characteristics based on entropy, skewness, and kurtosis, as well as intensity features based on ripplet transform. According to the findings, the updated MSER algorithm can handle pap smear images with worse quality due to inadequate straining and can also remove undesirable structures in the cell.

There were few methods that employed multi-level approaches for segmentation. Like, Zhang et al.<sup>17</sup> used a graph cut method integrated with textual and intensity-based features for segmentation. It was observed that all such methods rely on multi-level segmentation coupled with some post and pre-processing steps. Hence the failure at any level will affect the performance of the segmentation model, which in turn will also have a great effect on the classification accuracy and will increase the diagnosis error. Lu et al.<sup>18</sup> also implemented such a multi-level approach for segmentation, but their method failed on abnormal cells. This might be due to incomplete hand-crafted feature sets preventing the techniques from describing low-level features. However, hand-crafted features do not contain all the structural information of the nuclei; hence they result in poor segmentation performance. To enhance the segmentation performance, nuclei-type-specific criterion values should be used for the segmentation of different types of cervical nuclei with some post and pre-processing. This increases the length of the pipeline, and an error in any step will be convex to the error and will reflect on the subsequent steps.

The disadvantages discussed above can be addressed using deep learning methods. These methods have shown enhanced

performance in medical image segmentation, classification, lesion detection<sup>19,20</sup>. There are few studies that reported enhanced segmentation performance in terms of accuracy and efficiency while using DL methods. Zhao et al.<sup>21</sup> proposed a convolutional neural network-based deformable multipatch ensemble model for single-cell nuclei segmentation on the Herlev dataset. Liu et al.<sup>22</sup> built a segmentation model for single-cell nuclei segmentation by altering the structure of Mask R-CNN and adding fully linked conditional random fields. Lin et al.<sup>23</sup> used morphological convolutional neural networks to conduct multi-class and binary classification of single pap smear pictures. Song et al.<sup>24,25</sup> proposed a two-step approach, where in the first step, a deep learning method was used to segment the nuclei, and in the next step, a graph partitioning and superpixel approaches are used for the coarse-to-fine segmentation of the nuclei. A similar two-step approach was proposed by Zhang et al.<sup>26</sup>, where the authors segmented the single-cell nuclei by integrating convolutional neural networks (CNN) with graph-based approaches. Gautam et al.<sup>27</sup> developed a CNN model using transfer learning for single-cell nuclei segmentation.

With this motivation, a fully automatic cervical nuclei segmentation and classification approach was proposed in this paper. The proposed approach consists of a deep learning-based segmentation model, a fusion-based feature extraction model and a classification model. The structure of proposed work is shown in Figure 1. The contributions of the paper are summarised as follows:

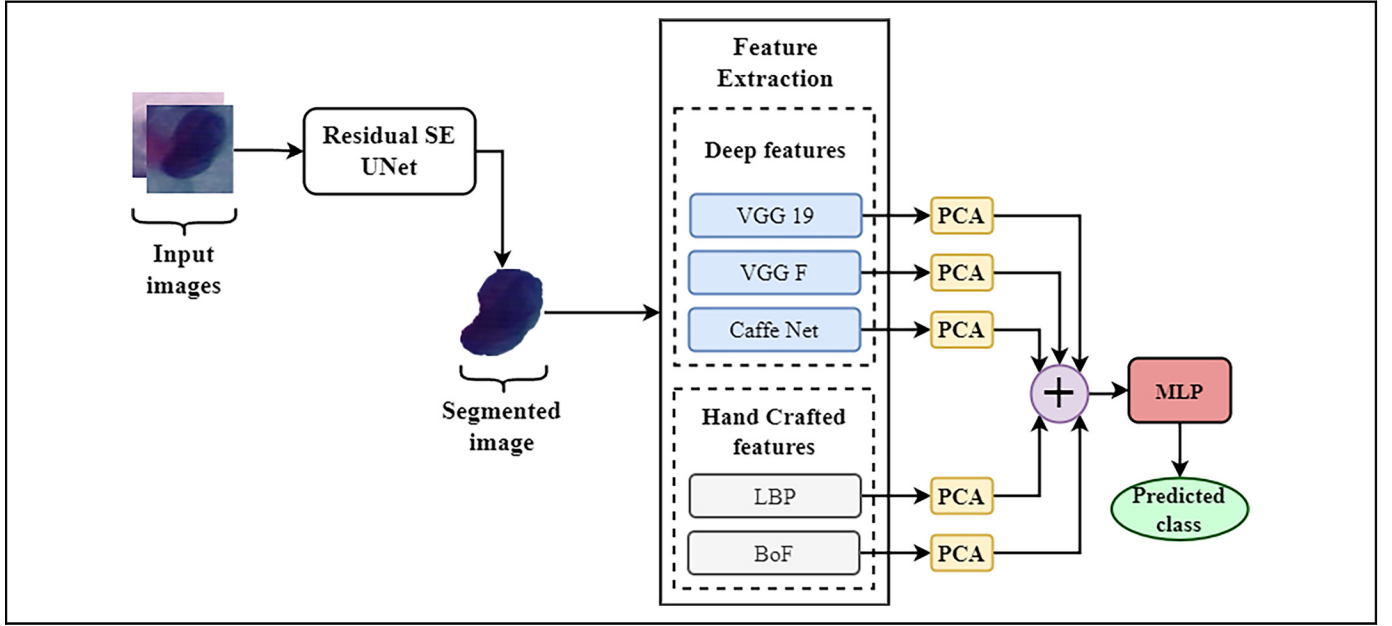
- The proposed work first segments the data, and then utilizes the segmented data for classification.
- The segmentation model is designed by modifying the structure of the U-Net model. A residual block with the Squeeze and Excitation (SE) block is used in place of the convolutional layers in each stage of the U-Net encoder-decoder network. This segmentation model was used to segment the nuclei from the cells.
- From the segmented image, the deep features and hand-crafted features are extracted and fused using the standard concatenation. To remove the redundancy among the extracted features, the PCA method is used before concatenation for feature selection.
- These fused features are used for training the multi-layer perceptron for classification.

The rest of the paper is organized as follows; Section 2 describes the materials and methods used, Section 3 describes experimental results and data used, section 4 presents the discussion, and section 5 concludes the work.

## Materials and Methods

### Residual SE UNet

In this work, a novel segmentation architecture based on UNet was proposed for the segmentation of nuclei from the pap smear images. The structure of the proposed network is shown in Figure 2. In this network, a residual block with the Squeeze



**Figure 1.** The structure of the proposed work.

and Excitation (SE) block is used in place of the convolutional layers in each stage of the UNet encoder-decoder network. This Residual SE module is shown in Figure 3. The residual block consists of a stack of two  $3 \times 3$  convolutional, two batch normalizations, 2 ReLU layers, and an auxiliary connection between the input and the output layer. The residual unit eases the network's learning, and the skip connection between the high and low levels in the residual unit will help for the propagation of information without degradation<sup>28</sup>.

In segmentation, spatial information is essential to identify the suspicious regions in the images. So, to improve the ability of the network to distinguish between the local and global information and to enhance its learning ability in each stage, an SE block is used after the residual block in this network. The SE block recalibrates the extracted features in two stages; in the first stage, the squeeze operation is performed where the features are globalized channel-wise into a one-dimensional (1D) array. In the second stage, these features are passed through two dense layers, and activations were described as weights to the input channels; this is referred to as excitation operation. These channel weights scale with the input features and enhance the feature representation ability of the network.

The network proposed in this work is a 9 level architecture consisting of three parts, namely encoder, decoder, and a bridge. The encoder converts the input image into a compact representation, and the decoder recovers this representation into pixel-wise classification. The bridge acts as a connection between encoder and decoder parts. The encoding block consists of four residual SE modules; specifically, the encoding block uses four downsampling operations after each Residual SE module to extract high-level semantic information. In each Residual SE encoding module, a stride of 2 is applied to the

first convolutional layer of the module to downsample the feature map by its half instead of using a pool operation to preserve positional information. Correspondingly, the decoder path consists of four Residual SE modules. There is a concatenation of the feature map from the corresponding encoding path with the upsampled feature map from the previous module. After the last encoding module, there is a  $1 \times 1$  convolutional layer and a sigmoid activation layer to project the desired segmented image.

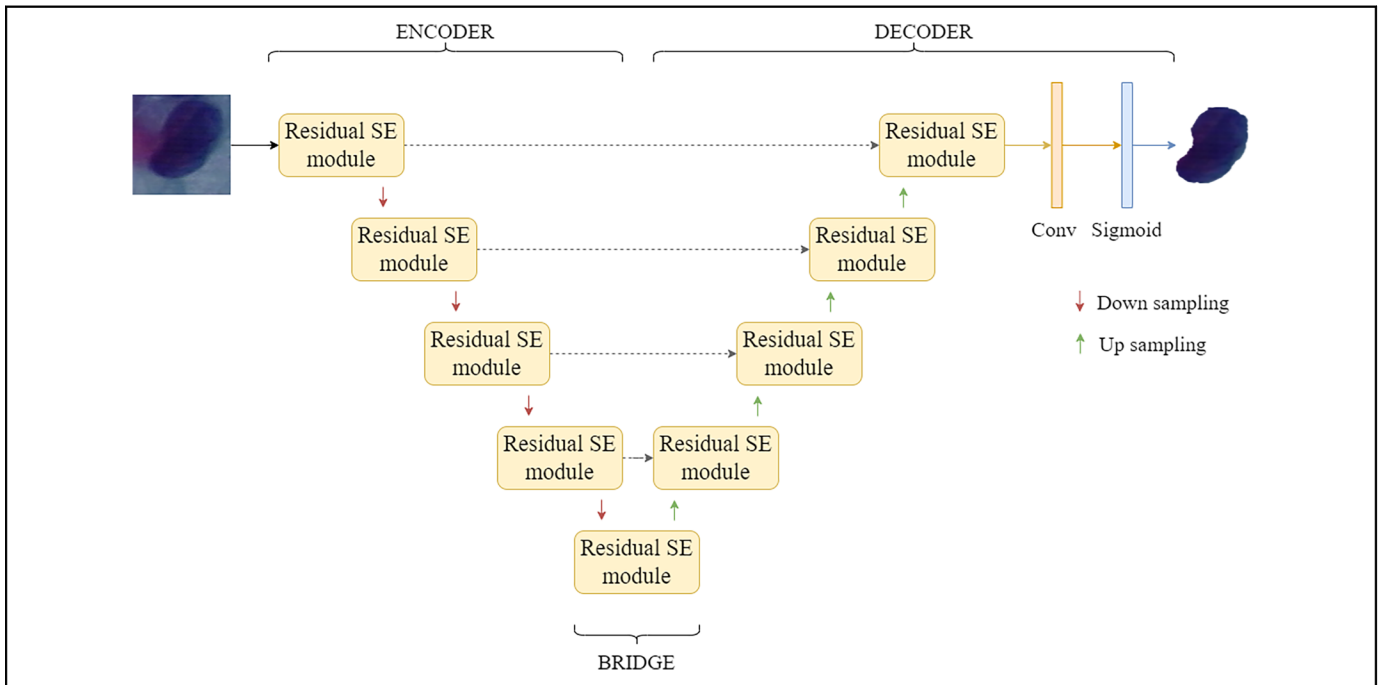
In the segmentation task, the imbalance between the background and the nucleus may result in segmentation bias. To deal with this problem, a loss function based on the dice coefficient is employed in this work. This loss function is presented in equation 1.

$$Loss_{seg} = 1 - \frac{1 + 2 \times Y_{seg} \times G_{seg}}{1 + Y_{seg} + G_{seg}} \quad (1)$$

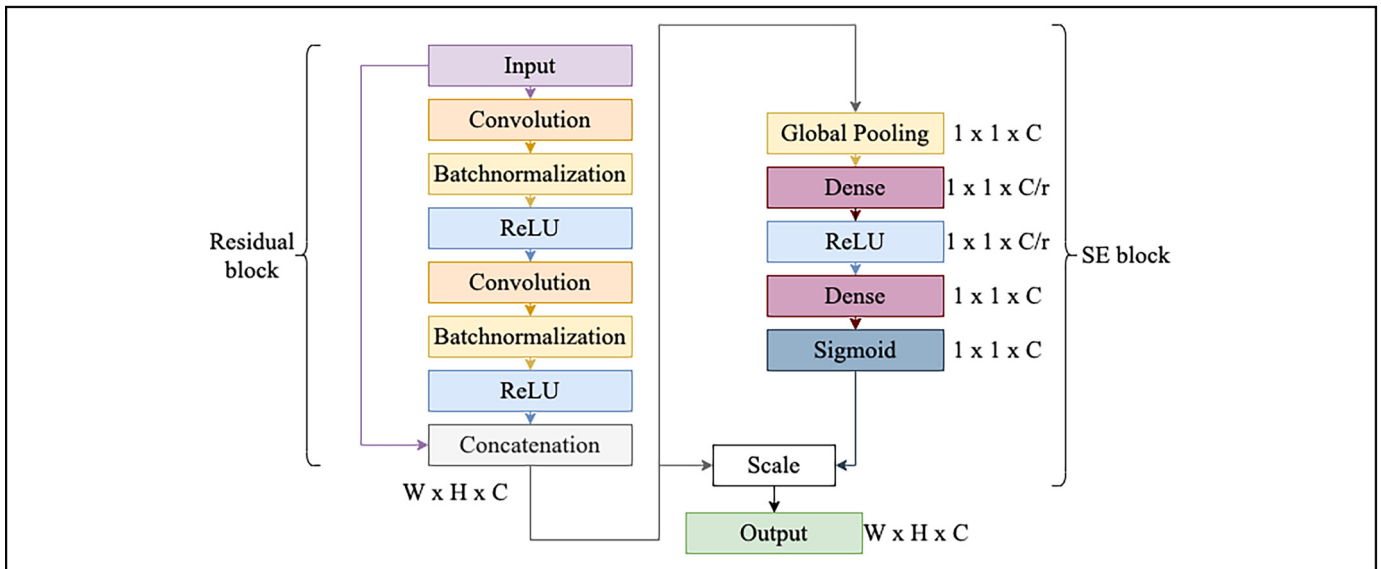
Where  $Y_{seg}$  represents the predicted mask by the Residual SE UNet,  $G_{seg}$  is the ground truth, and  $Loss_{seg}$  is the segmentation loss. This model is compiled using the Adam optimizer<sup>29</sup> with a batch size of 16, and the learning rate is set to  $1 \times 10^{-5}$ . The model is trained for 300 epochs with 100 steps for each epoch.

### Proposed classification model

The proposed multi-feature fusion approach consists of four main parts: (1) fine-tuning the pre-trained models to extract deep features, (2) computing LBP and BoF features for each segmented image, (3) Reducing the dimensions of the extracted features sets using PCA and (4) concatenating the hand-crafted and deep features for training the MLP for classification.



**Figure 2.** The structure of the proposed Residual SE UNet.



**Figure 3.** The structure of the proposed Residual Squeeze and Excitation module.

**Deep feature extraction.** In this work, three deep convolutional neural networks (DCNN), namely VGG19<sup>30</sup>, VGG-F<sup>31</sup>, and CaffeNet<sup>32</sup>, are employed for feature extraction. The VGG19 contains 16 convolutional layers with  $3 \times 3$  filters, five pooling layers, and three dense layers with 4096, 4096, and 1000 neurons. CaffeNet is a variant of AlexNet<sup>33</sup>; it contains five convolutional layers, three pooling layers, and three dense layers with the same number of neurons as the VGG19. VGG-F includes the same number of convolutional, pooling,

and dense layers as CaffeNet but with different filters. All these models are pre-trained on a natural dataset known ImageNet, which contains 14 million images categorized into 1000 classes. These models take an image of  $224 \times 224 \times 3$  dimensions and generate a prediction vector with 1000 dimensions.

The number of neurons in the final dense layer is modified to the number of classes in the dataset to make these models acceptable for the categorization of pap smear pictures. Then

the segmented images with size  $224 \times 224 \times 3$  are fed into the networks for fine-tuning. The Adam optimizer is chosen with a batch size of 32 and trained for 250 epochs. Since the training set used in this work is much smaller than the ImageNet, the learning rate is set to 0.0001, and this continues to drop by one-tenth for every 25 epochs, which prevents the models from overfitting. And 20 percent of the training images are randomly chosen as the validation set to evaluate the model during each epoch; if the error rate on the 80% training set continues to decrease and the error rate on the validation set stops declining, the training process is terminated even before reaching the maximum epoch limit. Then the 4096-dimensional output from the second last dense layers of the fine-tuned models is extracted and used for further classification.

**Hand-crafted feature.** In this work, Linear Binary Patterns (LBP)<sup>34</sup> and Bag of Features (BoF)<sup>35</sup> descriptors are used to characterize each image. The LBP descriptor is computed in three steps. For each pixel, the values of its eight neighboring pixels are assigned a binary value based on the center pixel; the value of the neighboring pixel is one of the existing values that is greater than the center pixel and is given zero if it is less than the center pixel. In the next step, these eight binary values are concatenated to form an eight-bit integer taking values from 0, . . . , 255. This process is carried out for all the pixels in the images. In the last step, the histogram of the frequency of each integer from the entire image is considered to be the 256 dimensional (D) descriptor. The computation of BoF also follows a three-step process. The Speed-Up-Robust-Feature (SURF)<sup>36</sup> approach is used to extract the key points and descriptors  $F = f_1, f_2, \dots, f_n$  from the images in the first stage, where each descriptor is a 128-dimensional vector. In the next phase, Vector Quantization (VQ) is used to assign descriptors F to the KBF clusters, also known as visual vocabulary. In the last step, the distribution of the SURF descriptors F over the visual vocabulary is counted as the BoF descriptor.

**Principal Component Analysis.** For each input image, there are three sets of deep features and two sets of hand-crafted features extracted, each of which has dimensions between 256 to 4096. The PCA algorithm is used on a group-by-group basis to deal with the curse of dimensionality and to select the most discriminative features. Let the  $j^{\text{th}}$  set of features be represented as  $Fe_{D,N}^j, D$ , where  $N$  represents the number of samples and  $D$  represents the dimension of the features. Then the single-value-decomposition (SVD)<sup>37</sup> is applied to the covariance matrix of  $Fe_{D,N}^j$  and its eigenvalues  $e_1, e_2, \dots, e_n$  are obtained. These eigenvalues and their accompanying eigenvectors  $v_1, v_2, \dots, v_n$  are listed in decreasing order. The features are then projected into a lower-dimensional space that are spanned by the eigenvectors, where the sum of the associated eigenvalues is greater than  $p$  percent of the total eigenvalues. The dimension of the generated features is controlled by the  $p$  parameter.

**Multi layer perceptron (MLP).** MLP is a multi-layer feed-forward deep neural network with a non-linear mapping of inputs to outputs. MLP is made up of three layers: an input layer, a hidden layer, and an output layer, in which each node is connected with suitable weights to all the nodes in the following layer. For training, MLP employs the backpropagation algorithm, which works by modifying the weights at each node, this backpropagation technique lowers the error transmitted throughout the network. The error  $E_i(n)$  at the  $i^{\text{th}}$  output node for the  $d^{\text{th}}$  data point is computed using equation 2.

$$E_i(n) = g(n)_i - y(n)_i \quad (2)$$

Where  $y(n)_i$  represents the predicted output, and the  $g(n)_i$  represents the actual output. This error can be minimized by correcting the weights at each node, presented in equation 3, with new weights computed using equation 4.

$$\sigma(n) = \frac{1}{2} \sum_i [E_i^2(n)] \quad (3)$$

$$\Delta W_{ij}(n) = -\alpha \frac{\partial \sigma(n)}{\partial_i(n)} g_j(n) \quad (4)$$

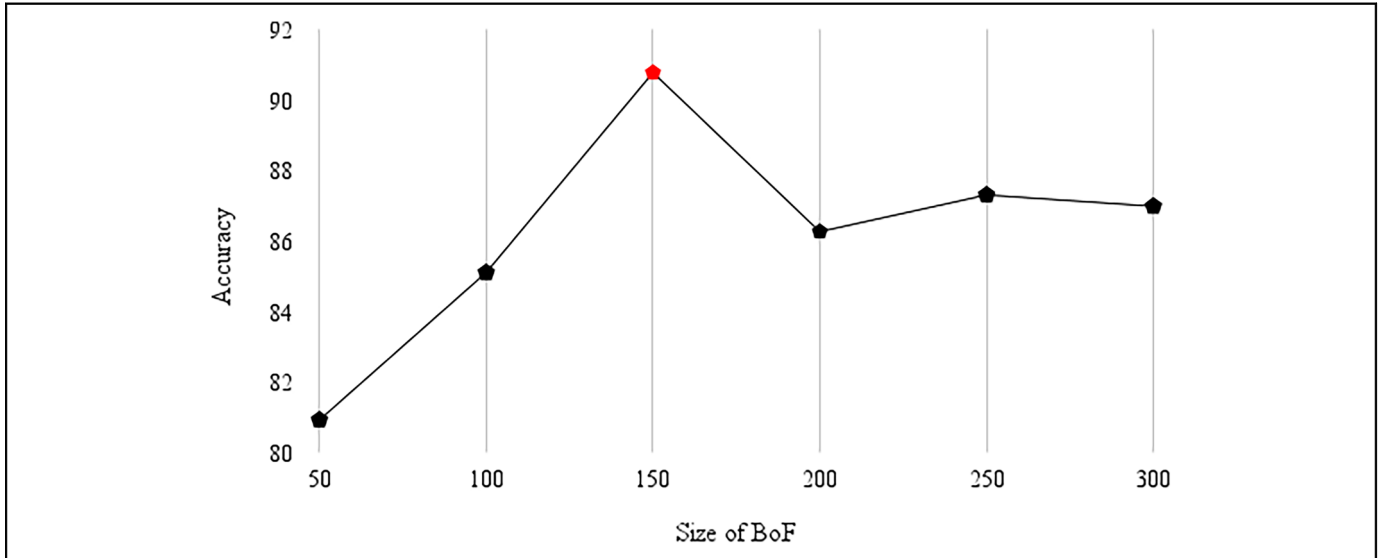
Where  $g_j(n)$  is the output of the previous node and  $\alpha$  is the learning rate. This process is repeated until the error becomes constant. This work uses one hidden layer by considering the advantages presented in<sup>38</sup> and is activated using the ReLU function.

**Parameter setting.** In this work, there are three types of parameters, namely DCNN based parameters, parameters related to dimensionality reduction, and parameters associated with MLP. Since we have opted for the transfer learning of the pre-trained DCNN's, only their weights and kernels are fine-tuned, but the structure and other parameters are unchanged. The parameters related to the dimensionality reduction, and MLP are set based on the performance reported on the validation test.

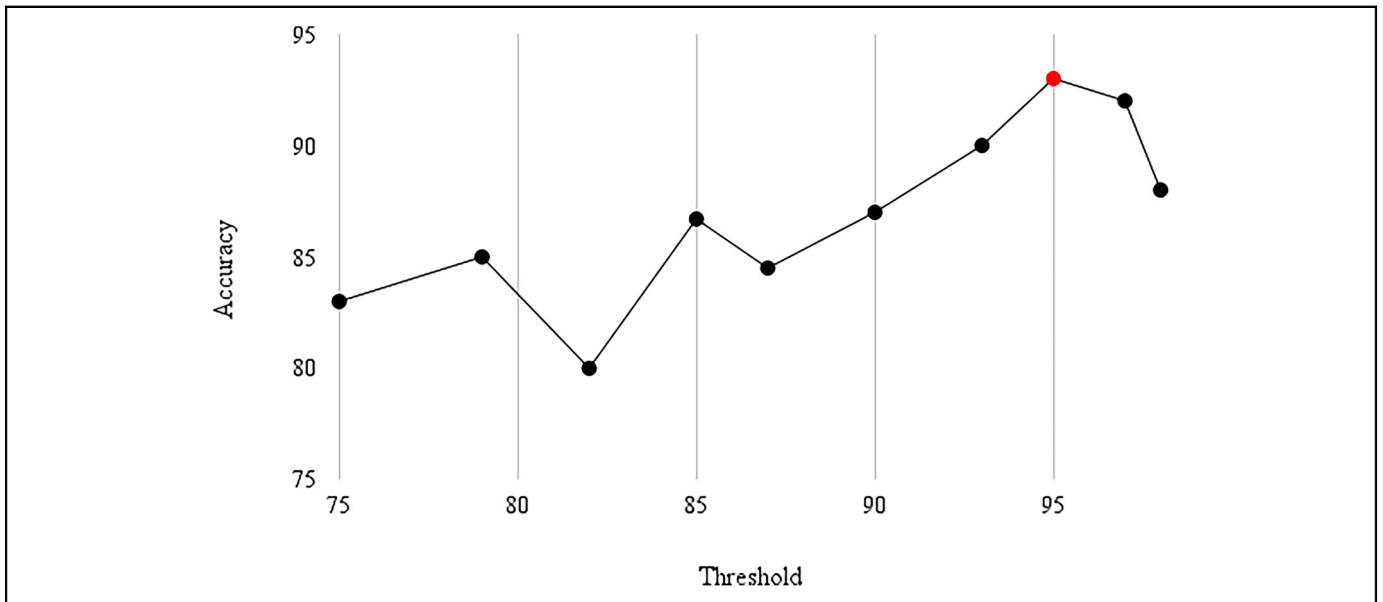
The dimensions of the BoF are based on the size of the visual vocabulary (KBF). Only the KBF is set to different values in this work, whereas the remaining parameters are unchanged. Figure 4 shows the accuracy graph reported by the proposed method on the validation set while varying the size of the KBF; the maximum accuracy is reported when the size of KBF is 150. So the size of the KBF is set as 150. (KBF = 150)

In PCA-based dimensionality reduction, the threshold value is changed, and other parameters are unchanged. Figure 5 shows the accuracy reported on the validation set by the proposed model with variant thresholds; the maximum accuracy was reported when the threshold was set to 95%.

In MLP, the batch size is changed by keeping the other parameters constant. The batch size is to different values starting from 8 to 64, and the graph of accuracy for each batch size is shown in Figure 6; the maximum accuracy was obtained on the validation set when the batch size is 32. So the batch size is set to 32.



**Figure 4.** Accuracy reported on the validation set for the variant sizes of BoF.



**Figure 5.** Accuracy reported on the validation set with threshold values.

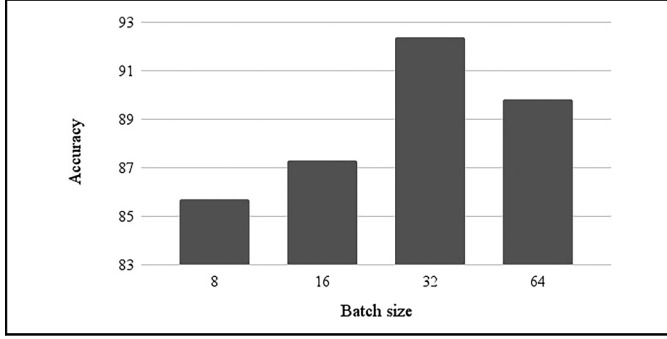
## Experimental Results

### Datasets

In this work, we employed three datasets namely Herlev, SIPaKMeD, and ISBI 2014 datasets for evaluation. Among these datasets the Herlev dataset is used for evaluating both segmentation, and classification models. Whereas the SIPaKMeD is used for evaluating the classification model, and the ISBI 2014 is used for evaluating the segmentation model.

*Herlev dataset.* is collected by the Herlev University Hospital using a microscope and digital camera<sup>39</sup>. The image resolution

used while acquiring the image is  $0.201 \mu\text{m}$  per pixel<sup>40</sup>. All the specimens are processed using the conventional pap straining and pap smear procedure. This Herlev dataset consists of 917 single cervical cell images classified into seven classes. An experienced doctor and two cyto-technicians manually annotate these images. The dataset also provides the ground truth images for training segmentation models. The categorical distribution of the dataset is shown in Table 1, and few sample images from the dataset are shown in Figure 7. As shown in Figure 7, most of the normal cells have smaller nuclei than the abnormal cells. And the SCCIS cells had a similar nucleus size as the CE cells, making the classification task



**Figure 6.** Accuracy reported on the validation set for different batch sizes.

**Table 1.** Class Distribution in Herlev Dataset.

Category	Class	No of Images
Abnormal	Squamous cell carcinoma in situ intermediate (SCCIS)	150
	Severe squamous non-keratinizing dysplasia (SSNKD)	197
	Moderate squamous non-keratinizing dysplasia (MSNKD)	146
	Mild squamous non-keratinizing dysplasia (MiSNKD)	182
Normal	Columnar epithelial (CE)	98
	Intermediate squamous epithelial (ISE)	70
	Superficial squamous epithelial (SQE)	74

challenging. In this work, the Herlev dataset is split into training, validation and testing sets based on the train-test strategy presented in<sup>41</sup>. Among the 917 images the 70% of images from each class are used for training, 10% are used for validation, and the rest 20% are used for testing the model.

*SIPaKMeD dataset*<sup>42</sup>. consists of 4049 isolated cervical cell images, which are manually cropped from 966 cluster cell images of Pap smear slides. The images are captured using a CCD camera adapted to an optical microscope. These cells are divided into five different classes. This class distribution is tabulated in Table 2. Among these the 60% of images from each class are used for training, 20% are used for validation, and the rest 20% are used for testing the model.

*ISBI 2014 dataset* is provided as a part of the Overlapping Cervical Cytology Image Segmentation Challenge ISBI 2014. This dataset contains 16 real images, and 945 synthetic images. The real images are of  $1024 \times 1024$  resolutions in grey-scale orientation. The real images are cropped into  $512 \times 512$  resolution, which were later enhanced into 1780 images. Among the 1780 images, the 1650 images are provided for training, and the rest 130 images are provided for testing the models. The 20% of the train set is used for validation.

## Performance metrics

**Segmentation metrics.** The Residual SE UNet is evaluated using pixel-based recall and precision measures. These measures are formulated in equations 5 and 6.

$$Recall_s = \frac{TP_s}{TP_s + FN_s} \quad (5)$$

$$Precision_s = \frac{TP_s}{TP_s + FP_s} \quad (6)$$

In the above equations 5 and 6,  $TP_s$  represents the number of pixels that are correctly predicted as the nuclei region,  $FP_s$  and  $FN_s$  represent the number of pixels that are wrongly predicted as background and nuclei regions. In addition to these measures Zijdenbos Similarity Index (ZSI) is also used for evaluation and is presented in equation 7.

$$ZSI = \frac{2 \times TP_s}{2 \times TP_s + FP_s + FN_s} \quad (7)$$

According to<sup>43</sup>, the predicted mask and ground truth are excellently matched when ZSI is greater than 0.7.

**Classification metrics.** The classification network is evaluated using accuracy, recall, specificity, precision, and F1-score. The metrics can be calculated using equations 8-12.

$$Accuracy_c = \frac{\sum_{i=1}^7 \text{No of correctly classified samples as class } i}{\sum_{i=1}^7 \text{Total number of samples in class } i} \quad (8)$$

$$Recall_c = \frac{TP_c}{TP_c + FN_c} \quad (9)$$

$$Specificity_c = \frac{TN_c}{TN_c + FP_c} \quad (10)$$

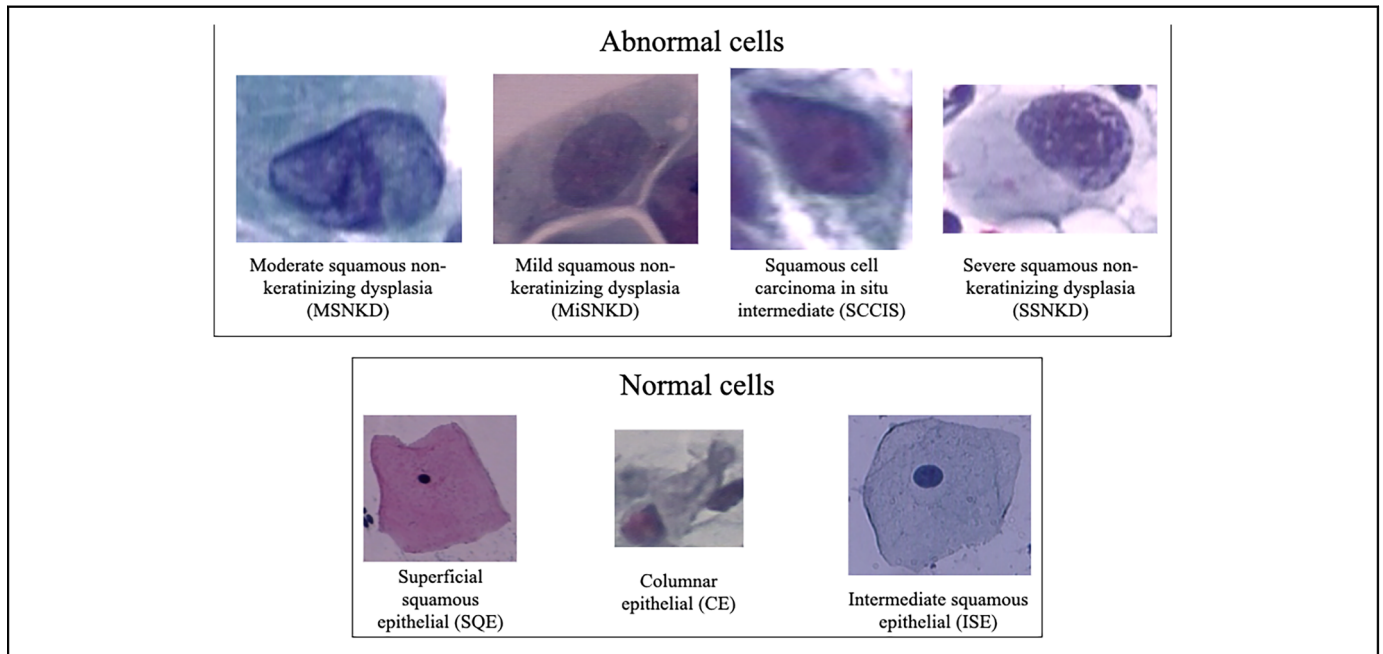
$$Precision_c = \frac{TP_c}{TP_c + FP_c} \quad (11)$$

$$F1 - score_c = \frac{2 \times Precision_c \times Recall_c}{Precision_c + Recall_c} \quad (12)$$

In the equations 5-12,  $TP_c$  represents the number of images correctly classified as abnormal,  $TN_c$  represents the number of images that are correctly classified as normal,  $FP_c$  and  $FN_c$  represents the number of images that are wrongly classified as abnormal and normal.

## Ablation study

**Residual SE UNet.** We have also performed an ablation study to understand the efficiency of each module in Residual SE UNet. These ablation experiments are performed on the Herlev dataset. The results of this study are shown in Table 3. From



**Figure 7.** Sample images from the Herlev dataset.

**Table 2.** Class Distribution In SIPaKMeD Dataset .

Category	Class	No of Images
Normal	Parabasal (PARA)	787
	Superficial-intermediate (SI)	831
Abnormal	Dyskeratotic (DYSK)	813
	Koilocytotic (KOIL)	825
Benign	Metaplastic (META)	793

**Table 3.** Abilation Study of the Residual SE UNet.

Models	Precision	Recall	ZSI
Standard UNet	86.59	89.05	0.83
UNet + Residual module (without SE block)	90.19	92.31	0.90
<b>UNet + Residual module (with SE block)</b>	<b>97.24</b>	<b>96.2</b>	<b>0.97</b>

Table 3, it can be seen that there is an performance enhancement with the addition of Residual SE Modules to the standard UNet.

**Classification network.** In this work, the performance of the Feature Concatenation Approach was assessed using three groups of feature representations. The first group represents the performance reported by the proposed approach while using sole hand-crafted features, the second group represents the performance reported by combining hand-crafted features with the deep features extracted by fully trained models, and the third group presents the performance reported while using hand-crafted features with deep features extracted by the fine-tuned and pre-trained models.

From Table 4, it was observed that the feature representations learned by the transfer learning model reported better classification accuracy than the hand-crafted features. However, the fully-trained models reported worse accuracy than the hand-crafted features. The features extracted by the fully-trained models, when used solely or when combined with the hand-crafted features, also reported worse accuracy than the pre-trained transfer learning models. This shows the advantage of using the transfer learning approach while having data scarcity

and other constraints. In addition, the concatenation of deep and hand-crafted features reported significantly better classification accuracy.

## Results reported

### Segmentation

The proposed Residual SE UNet is evaluated using the Herlev, and the ISBI 2014 datasets. The Table 5 presents the precision, recall, and ZSI scores reported by the proposed model for segmenting the 7 types of cervical nuclei on Herlev dataset. Furthermore, the average results of 7 types achieved by the Residual SE UNet on both the datasets is compared with the other existing approaches. These results are shown in Tables 6 and 7. The proposed segmentation model reported a precision of 97.24%, recall of 96.2%, and ZSI of 0.97 on the test set.

The proposed model reported better average precision, and ZSI than the existing works on Herlev dataset. Also, the proposed model reported better average precision and recall than existing works on ISBI 2014 dataset. Figure 8 shows the



**Table 4.** Accuracy Reported While Using Different Feature Sets.

Methods	Features Sets	Accuracy
Hand crafted features	BoF	75.98
	LBP	78.43
Deep features extracted by fully training CNN's	Deep features from VGG19	72.63
	Deep features from VGG-F	71.32
	Deep features from Caffe Net	69.50
	3 Deep features	74.92
Deep features extracted by pre-trained CNN's	Deep features from pre-trained VGG19	90.19
	Deep feature from pre-trained VGG-F	91.12
	Deep feature from pre-trained Caffe Net	89.42
	3 Deep features	93.00
Deep features extracted by fully training CNN's with hand-crafted features	3 Deep features + LBP	81.05
	3 Deep features + BoF	82.78
	3 Deep features + LBP + BoF	85.12
	3 Deep features from pre-trained models + LBP	93.13
Deep features extracted by pre-trained CNN's through fine tuning with hand-crafted features	3 Deep features from pre-trained models + BoF	95.45
	3 Deep features from pre-trained models + BoF + LBP (proposed)	<b>98.39</b>

qualitative comparison of the segmented output reported by the proposed model and the existing methods on images from the Herlev dataset. It can be observed that the proposed segmentation model can generate accurate nucleus boundaries for a wide variety of nuclei with irregular nuclei shape, size, and non-uniform chromatin distribution.

### Classification

In this work, we evaluated the proposed classification approach by using two datasets namely Herlev, and SIPaKMeD. In medical informatics, recall is considered to be the most important metric<sup>48,49</sup>. Table 8 presents the recall reported for each class while using hand-crafted features, deep features extracted by transfer learning models, and jointly using both (proposed feature concatenation). The highest recall is highlighted in bold. The proposed feature concatenation approach performed best in 5 out of 7 categories and is also higher than the other feature sets on both the datasets.

The proposed feature concatenation approach is also compared with the existing methods<sup>39,50,14,12,51-54</sup>. The above methods are downloaded from their public implementations

**Table 5.** Comparison of Class Specific Precision, Recall, and ZSI Reported by the Residual SE UNet with the Existing Works Employing Herlev Dataset.

Methods	Class	Precision	Recall	ZSI	
Multi-scale hierarchical segmentation algorithm <sup>15</sup>	SSNKD	90.12	89.39	0.921	
	SQE	69.37	63.48	0.848	
	ISE	79.29	73.31	0.914	
	CE	85.15	77.58	0.892	
	MSNKD	91.00	86.78	0.904	
	MiSNKD	88.64	86.73	0.895	
	SCCIS	90.35	89.36	0.913	
	Mean	84.84	80.94	0.898	
	Mask RCNN +LFC +CRF <sup>22</sup>	SSNKD	96.06	97.12	0.951
		SQE	95.05	97	0.950
ISE		93.10	94.17	0.921	
CE		93.09	94.51	0.821	
MSNKD		96.04	98.17	0.97	
MiSNKD		96.04	97.08	0.96	
SCCIS		97.04	95.12	0.951	
Mean		95.20	96.16	0.932	
Radiating Gradient Vector flow <sup>13</sup>		SSNKD	88.23	89.66	0.879
		SQE	92.05	88.10	0.898
	ISE	95.12	92.42	0.869	
	CE	86.79	76.57	0.821	
	MSNKD	89.27	86.84	0.875	
	MiSNKD	92.29	90.44	0.862	
	SCCIS	84.17	90.88	0.867	
	Mean	89.70	87.84	0.867	
	Proposed Residual SE UNet	SSNKD	95.68	94.31	0.950
		SQE	98.71	98.29	0.989
ISE		97.78	95.04	0.995	
CE		98.84	95.31	0.942	
MSNKD		96.18	95.69	0.977	
MiSNKD		97.94	98.00	0.969	
SCCIS		95.59	96.80	0.973	
Mean		<b>97.24</b>	<b>96.20</b>	<b>0.970</b>	

**Table 6.** Comparison of Average Precision, Recall, and ZSI Reported by the Residual SE UNet with the Existing Works Employing Herlev Dataset.

Methods	Precision	Recall	ZSI
Multi-scale hierarchical segmentation algorithm <sup>15</sup>	84.84	80.94	0.898
Mask RCNN + LFC + CRF <sup>22</sup>	96.00	96.00	0.950
Radiating Gradient Vector flow <sup>13</sup>	89.70	87.84	0.867
D-MEM (U-Net) <sup>21</sup>	94.60	<b>98.40</b>	0.933
PGU-net+ <sup>44</sup>	90.10	96.80	0.926
Proposed Residual SE UNet	<b>97.24</b>	96.20	<b>0.970</b>

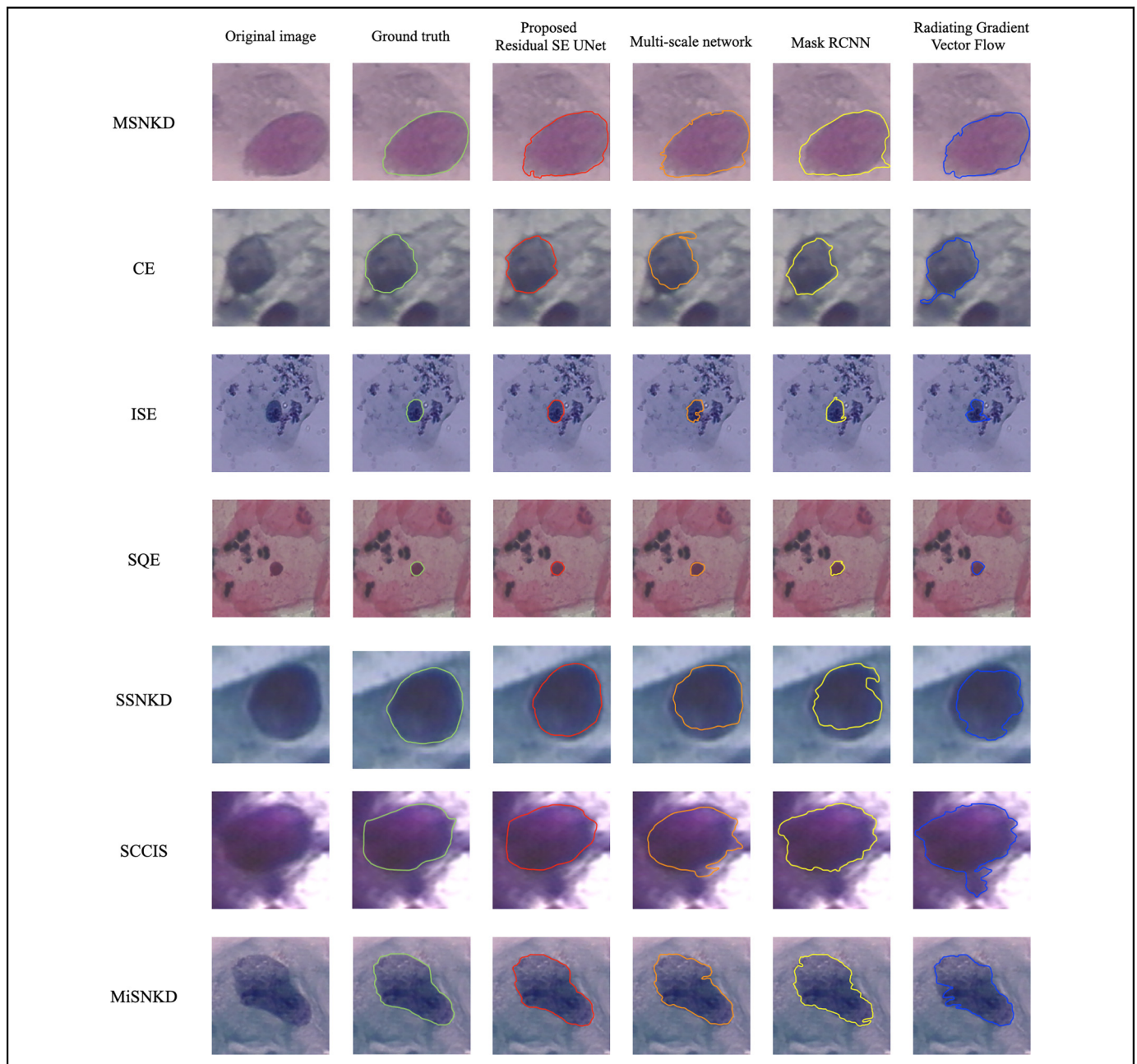
and trained and tested with the same evaluation protocol used by the proposed work on the Herlev, and SIPAKMED datasets for a fair comparison. This comparison in terms of accuracy, recall, and specificity is shown in Table 9. The proposed model reported higher accuracy than the existing works.

**Table 7.** Comparison of Precision, and Recall Reported by the Residual SE UNet with the Existing Works Employing ISBI 2014 Dataset.

Method	Precision	Recall
HMLS <sup>45</sup>	93.81	92.34
MSERLS <sup>18</sup>	94.23	91.82
LTSN <sup>46</sup>	97.64	94.00
DeepCNN1 <sup>47</sup>	94.61	95.59
<b>Proposed model</b>	<b>98.32</b>	<b>97.18</b>

### Computational complexity

The proposed work is implemented in Pycharm. The overall processing of a pap-smear image of  $224 \times 224 \times 3$  pixel resolution on a PC with 2.4 GHz dual-core Intel i5 and 4 GB took 21s. These running times were reported on the images from the single-cell Herlev dataset. In this work, code profile analysis is also performed to understand the time consumed by the segmentation and classification method. Among the two methods, the segmentation method consumed 12s, and the classification



**Figure 8.** Qualitative comparison of the proposed model with existing methods. The ground truth boundary is indicated by green color, the boundary predicted by the Residual SE UNet is indicated by red color, the boundary predicted by the Multi-scale hierarchical segmentation algorithm<sup>15</sup>, Mask RCNN + LFC + CRF<sup>22</sup>, Radiating Gradient Vector flow<sup>13</sup> are indicated by orange, yellow and blue colours.

**Table 8.** Recall Reported for Each Class While using Different Transfer Learning Models.

Dataset	Classes	VGG19	VGG-F	Caffe Net	LBP	BoF	Proposed Model
Herlev	SSNKD	90.39	89.78	93.10	83.60	86.19	<b>99.12</b>
	SQE	<b>99.36</b>	94.24	95.63	79.53	82.92	98.76
	ISE	91.21	92.08	94.42	88.27	89.16	99.45
	CE	95.68	89.26	93.42	87.31	89	<b>99.39</b>
	MSNKD	93.37	96.24	95.19	82.36	86.29	<b>98.92</b>
	MiSNKD	96.30	<b>99.37</b>	94.53	75.29	70.50	98.42
	SCCIS	94.59	95.91	94.25	84.45	80.26	<b>98.74</b>
	Mean	94.41	93.84	94.36	82.97	83.47	<b>98.97</b>
SIPAKMED	PARA	97.16	98.33	98.62	90.42	88.53	<b>98.72</b>
	SI	96.72	99.41	98.89	85.89	83.29	<b>99.62</b>
	DYSK	97.11	97.52	97.81	87.14	86.31	<b>98.54</b>
	KOIL	96.39	98.67	97.89	91.55	85.68	<b>99.25</b>
	META	96.11	99.01	98.87	89.33	84.52	<b>99.62</b>
	Mean	96.69	98.59	98.41	88.86	85.66	<b>99.15</b>

**Table 9.** Comparison of Accuracy, Recall, and Specificity Reported by the Proposed Model with Other Existing Works on the Herlev, and SIPAKMED Datasets.

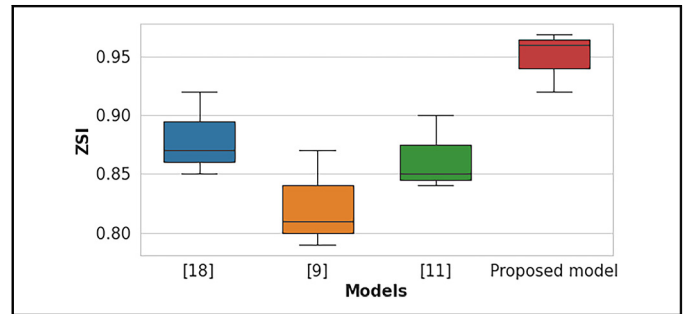
Dataset	Methods	Accuracy	Recall	Specificity
Herlev	DeepCervix <sup>55</sup>	90.30	91.10	–
	Jantzen et al. <sup>39</sup>	93.60	97.50	85.60
	Marinakis et al. <sup>50</sup>	96.70	98.40	92.20
	Marinakis et al. <sup>14</sup>	96.80	98.50	92.10
	Chankong et al. <sup>12</sup>	97.80	98.30	96.50
	Liu et al. <sup>56</sup>	92.35	93.50	–
	<b>Proposed model</b>	<b>98.39</b>	<b>98.97</b>	<b>97.65</b>
	SIPAKMED	DeepPap <sup>52</sup>	93.58	97.40
Win et al. <sup>53</sup>		94.09	–	–
CompactVGG <sup>51</sup>		97.80	98.30	99.17
Qin et al. <sup>54</sup>		98.14	98.10	99.53
DeepCervix <sup>55</sup>		99.14	99.00	–
<b>Proposed model</b>		<b>99.16</b>	<b>99.15</b>	<b>99.75</b>

method consumed 9s (including 7s for feature extraction and 2s for classification).

## Discussion

The experimental findings show that the proposed models can accurately segment and classify cervical nuclei from pap smear images with good precision, recall, specificity, and ZSI. The following are the primary points that emphasise the proposed work.

- The manual morphological analysis of cellular images for diagnosing cervical cancer from pap smear slides on a large scale is a time-consuming and tedious task. And a manual examination of these slides often contains human error<sup>57,58,3</sup>, resulting in false-positive/negative findings. Automated segmentation and classification of the nuclei will help rapidly assess pap smear slides on a large scale with zero human error and less diagnostic

**Figure 9.** Comparison of ZSI values reported by the Residual SE UNet with existing methods on Herlev dataset.

time than the manual procedure. This work is advantageous as it can segment and classify cervical nuclei with high accuracy, precision, recall, and ZSI, enabling rapid nuclear-quantification analysis.

- Even though the proposed segmentation method is computationally more expensive than the multi-scale network<sup>15</sup> (64 vs 59) and Mask R-CNN<sup>22</sup> (64 vs 62) methods, this can be optimized by using sophisticated hardware.
- A box plot is presented in Figure 9, which shows the distribution of the ZSI metric reported by the proposed segmentation model, multi-scale network<sup>15</sup>, MaskRCNN<sup>22</sup>, and Radiating Gradient Vector flow<sup>13</sup> on the Herlev test set. The proposed segmentation method has a higher median of ZSI than the other three methods. This demonstrates the proposed segmentation model's superiority over existing techniques.
- The proposed approach does not involve any pipeline methods and pre-processing methods discussed in the literature. It directly takes the pap-smear image as input and segment cervical nuclei. The features extracted from the segmented cervical nuclei and used for the classification. The proposed segmentation and classification models reported higher performance (in terms of

precision, recall and ZSI for segmentation task and accuracy, recall, and specificity for classification task) than the existing works that employed pre and post-processing methods (reported in Tables 6 and 9).

- Even though our proposed segmentation and classification models enhanced the performance in segmenting and classifying cervical cytopathology cell images, our models have the following limitations. The performance of our algorithms needs further perfection for real pre-clinical use. Moreover, we have not explored the possibility of any data resampling for balancing the dataset that may result in better performance.

## Conclusion

This work proposes two deep learning-based approaches for the segmentation and classification of cervical nuclei. The segmentation network was designed using the well-known architecture UNet as the backbone, and residual SE modules are designed for efficient feature extraction. These modules are used in place of the convolutional layers in the standard UNet for segmentation. From the segmented nuclei, three sets of deep features and two sets of hand-crafted features are extracted, and PCA is used to reduce these features' dimensions for concatenation. The single layer perceptron is employed for classification. These methods are trained and evaluated using the Herlev, SIPaKMeD, and ISBI 2014 datasets. Among these datasets the Herlev dataset is used for evaluating both segmentation, and classification models. Whereas the SIPaKMeD is used for evaluating the classification model, and the ISBI 2014 is used for evaluating the segmentation model. Both the segmentation and classification works reported better performance than the existing works in the literature. We anticipate that these methods help for the rapid diagnosis of cervical cancer in the early stage, thus reducing the mortality rate and helping patients for a faster diagnosis. In the future work, different pre-processing methods such as transfer learning, and vision-transformer-based approaches can be studied to diagnose cervical cancer from pap smear images.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the publication of this article: The authors are thankful for the financial support provided by the Intelligent Systems Research Centre (ISRC), Ulster University, UK.

## ORCID iD

Pratheepan Yogarajah  <https://orcid.org/0000-0002-4586-7228>

## References

1. Davey E, Barratt A, Irwig L, Chan SF, Macaskill P, Mannes P, Saville AM. Effect of Study Design and Quality on Unsatisfactory Rates, Cytology Classifications, and Accuracy in Liquid-based Versus Conventional Cervical Cytology: a Systematic Review. *Lancet*. 2006; 367(9505): 122-s 132.
2. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer. *Am J Clin Pathol*. 2012; 137(4): 516-542.
3. Birdsong GG. Automated Screening of Cervical Cytology Specimens. *Hum Pathol*. 1996; 27(5): 468-481.
4. Kitchener HC, Blanks R, Dunn G, Gunn L, Desai M, Albrow R, Mather J, Rana DN, Cubie H, Moore C, Legood R. Automation-assisted Versus Manual Reading of Cervical Cytology (MAVARIC): a Randomised Controlled Trial. *Lancet Oncol*. 2011; 12(1): 56-64.
5. Chowdary GJ. Machine learning and deep learning methods for building intelligent systems in medicine and drug discovery: A comprehensive survey. *arXiv preprint arXiv:2107.14037*. 2021 Jul 19.
6. Chowdary GJ. Class dependency based learning using Bi-LSTM coupled with the transfer learning of VGG16 for the diagnosis of Tuberculosis from chest x-rays. *arXiv preprint arXiv:2108.04329*. 2021 Jul 19.
7. Chowdary J, Yogarajah P, Chaurasia P, Guruviah V. A multi-task learning framework for automated segmentation and classification of breast tumors from ultrasound images. *Sage Ultrasonic Imaging*. 2022; 44(1): 3-12.
8. Bengtsson E, Malm P. Screening for cervical cancer using automated analysis of PAP-smears. *Comput Math Methods Med*. 2014; 1-12.
9. Zhang L, Kong H, Ting Chin C, Liu S, Fan X, Wang T, Chen S. Automation-assisted Cervical Cancer Screening in Manual Liquid-based Cytology with Hematoxylin and Eosin Staining. *Cytometry Part A*. 2014; 85(3): 214-230.
10. Rahaman MM, Li C, Wu X, Yao Y, Hu Z, Jiang T, Li X, Qi S. A Survey for Cervical Cytopathology Image Analysis Using Deep Learning. *IEEE Access*. 2020; 8: 61687-61710.
11. Dong N, Zhao L, Wu A. Cervical Cell Recognition Based on AGVF-Snake Algorithm. *Int J Comput Assist Radiol Surg*. 2019; 14(11): 2031-2041.
12. Chankong T, Theera-Umpun N, Auephanwiriyaikul S. Automatic Cervical Cell Segmentation and Classification in Pap Smears. *Comput Methods Programs Biomed*. 2014; 113(2): 539-556.
13. Li K, Lu Z, Liu W, Yin J. Cytoplasm and Nucleus Segmentation in Cervical Smear Images Using Radiating GVF Snake. *Pattern Recognit*. 2012; 45(4): 1255-1264.
14. Marinakis Y, Dounias G, Jantzen J. Pap Smear Diagnosis Using a Hybrid Intelligent Scheme Focusing on Genetic Algorithm Based Feature Selection and Nearest Neighbor Classification. *Comput Biol Med*. 2009; 39(1): 69-78.
15. Gençtav A, Aksoy S, Önder S. Unsupervised Segmentation and Classification of Cervical Cell Images. *Pattern Recognit*. 2012; 45(12): 4151-4168.
16. Bora K, Chowdhury M, Mahanta LB, Kundu MK, Das AK. Automated Classification of Pap Smear Images to Detect

- Cervical Dysplasia. *Comput Methods Programs Biomed.* 2017; 138: 31-47.
17. Zhang L, Kong H, Chin CT, Liu S, Chen Z, Wang T, Chen S. Segmentation of Cytoplasm and Nuclei of Abnormal Cells in Cervical Cytology Using Global and Local Graph Cuts. *Comput Med Imaging Graph.* 2014; 38(5): 369-380.
  18. Lu Z, Carneiro G, Bradley AP. An Improved Joint Optimization of Multiple Level Set Functions for the Segmentation of Overlapping Cervical Cells. *IEEE Trans Image Process.* 2015; 24(4): 1261-1272.
  19. Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A Survey on Deep Learning in Medical Image Analysis. *Med Image Anal.* 2017; 42: 60-88.
  20. LeCun Y, Bengio Y, Hinton G. Deep Learning. *Nature.* 2015; 521(7553): 436-444.
  21. Zhao J, Li Q, Li X, Li H, Zhang L. Automated segmentation of cervical nuclei in pap smear images using deformable multi-path ensemble model. In *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* 2019 Apr 8 (pp. 1514-1518). IEEE.
  22. Liu Y, Zhang P, Song Q, Li A, Zhang P, Gui Z. Automatic Segmentation of Cervical Nuclei Based on Deep Learning and a Conditional Random Field. *IEEE Access.* 2018; 6: 53709-53721.
  23. Lin H, Hu Y, Chen S, Yao J, Zhang L. Fine-grained Classification of Cervical Cells Using Morphological and Appearance Based Convolutional Neural Networks. *IEEE Access.* 2019; 7: 71541-71549.
  24. Song Y, Zhang L, Chen S, Ni D, Li B, Zhou Y, Lei B, Wang T. A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2014 Aug 26 (pp. 2903-2906). IEEE.
  25. Song Y, Zhang L, Chen S, Ni D, Lei B, Wang T. Accurate Segmentation of Cervical Cytoplasm and Nuclei Based on Multiscale Convolutional Network and Graph Partitioning. *IEEE Trans Biomed Eng.* 2015; 62(10): 2421-2433.
  26. Zhang L, Sonka M, Lu L, Summers RM, Yao J. Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei. In *IEEE 14th international symposium on biomedical imaging (ISBI 2017)* 2017 Apr 18 (pp. 406-409). IEEE.
  27. Gautam S, Bhavsar A, Sao AK, Harinarayan KK. CNN based segmentation of nuclei in PAP-smear images with selective pre-processing. In *Medical Imaging 2018: Digital Pathology* 2018 Mar 6 (Vol. 10581, pp. 246-254). SPIE.
  28. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* 2016 (pp. 770-778).
  29. Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.* 2014 Dec 22.
  30. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.* 2014 Sep 4.
  31. Chatfield K, Simonyan K, Vedaldi A, Zisserman A. Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531.* 2014 May 14.
  32. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia* 2014 Nov 3 (pp. 675-678).
  33. Krizhevsky A, Sutskever I, Hinton GE. Imagenet Classification with Deep Convolutional Neural Networks. *Adv Neural Inf Process Syst.* 2012; 25: 1097-1105.
  34. Ojala T, Pietikäinen M, Harwood D. A Comparative Study of Texture Measures with Classification Based on Featured Distributions. *Pattern Recognit.* 1996; 29(1): 51-59.
  35. Fei-Fei L, Perona P. A bayesian hierarchical model for learning natural scene categories. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* 2005 Jun 20 (Vol. 2, pp. 524-531). IEEE.
  36. Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up Robust Features (SURF). *Comput Vis Image Underst.* 2008; 110(3): 346-359.
  37. Jolliffe IT. Principal components in regression analysis. In *Principal component analysis* 1986 (pp. 129-155). Springer, New York, NY.
  38. Huang GB, Chen YQ, Babri HA. Classification Ability of Single Hidden Layer Feedforward Neural Networks. *IEEE Trans Neural Netw.* 2000; 11(3): 799-801.
  39. Jantzen J, Norup J, Dounias G, Bjerregaard B. Pap-smear benchmark data for pattern classification. *Nature inspired Smart Information Systems (NiSIS 2005).* 2005 Oct 3:1-9.
  40. Martin L, Exbrayat M. Pap-smear classification. 2003.
  41. Dobbin KK, Simon RM. Optimally Splitting Cases for Training and Testing High Dimensional Classifiers. *BMC Med Genomics.* 2011; 4(1): 1-8.
  42. Plissiti ME, Dimitrakopoulos P, Sfikas G, Nikou C, Krikoni O, Charchanti A. SIPAKMED: A new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images. In *25th IEEE International Conference on Image Processing (ICIP)* 2018 Oct 7 (pp. 3144-3148). IEEE.
  43. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. *IEEE Trans Med Imaging.* 1994; 13(4): 716-724.
  44. Zhao J, Dai L, Zhang M, Yu F, Li M, Li H, Wang W, Zhang L. PGU-net+: progressive growing of U-net+ for automated cervical nuclei segmentation. In *International Workshop on Multiscale Multimodal Medical Imaging* 2019 Oct 13 (pp. 51-58). Springer, Cham.
  45. Braga AM, Marques RC, Medeiros FN, Neto JF, Bianchi AG, Carneiro CM, Ushizima DM. Hierarchical Median Narrow Band for Level Set Segmentation of Cervical Cell Nuclei. *Measurement.* 2021; 176: 109232.
  46. Ushizima DM, Bianchi AG, Carneiro CM. Segmentation of sub-cellular compartments combining superpixel representation with voronoi diagrams. *Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States);* 2015 Jan 5.
  47. Tareef A, Song Y, Huang H, Wang Y, Feng D, Chen M, Cai W. Optimizing the Cervix Cytological Examination Based on Deep

- Learning and Dynamic Shape Modeling. *Neurocomputing*. 2017; 248: 28-40.
48. Hoi SC, Jin R, Zhu J, Lyu MR. Batch mode active learning and its application to medical image classification. In *Proceedings of the 23rd international conference on Machine learning 2006 Jun 25* (pp. 417-424).
49. Taha AA, Hanbury A. Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool. *BMC Med Imaging*. 2015; 15(1): 1-28.
50. Marinakis Y, Marinaki M, Dounias G. Particle Swarm Optimization for Pap-smear Diagnosis. *Expert Syst Appl*. 2008; 35(4): 1645-1656.
51. Chen H, Liu J, Wen QM, Zuo ZQ, Liu JS, Feng J, Pang BC, Xiao D. CytoBrain: Cervical Cancer Screening System Based on Deep Learning Technology. *J Comput Sci Technol*. 2021; 36(2): 347-360.
52. Zhang L, Lu L, Nogues I, Summers RM, Liu S, Yao J. DeepPap: Deep Convolutional Networks for Cervical Cell Classification. *IEEE J Biomed Health Inform*. 2017; 21(6): 1633-1643.
53. Win KP, Kitjaidure Y, Hamamoto K, Myo Aung T. Computer-assisted Screening for Cervical Cancer Using Digital Image Processing of Pap Smear Images. *Appl Sci*. 2020; 10(5): 1800.
54. Qin J, He Y, Ge J, Liang Y. A multi-task feature fusion model for cervical cell classification. *IEEE J Biomed Health Inform*. 2022; 26(9): 4668-4678.
55. Rahaman MM, Li C, Yao Y, Kulwa F, Wu X, Li X, Wang Q. DeepCervix: a Deep Learning-based Framework for the Classification of Cervical Cells Using Hybrid Deep Feature Fusion Techniques. *Comput Biol Med*. 2021; 136: 104649.
56. Liu W, Li C, Xu N, et al. CVM-Cervix: A hybrid cervical pap-smear image classification framework using CNN, visual transformer and multilayer perceptron. *Pattern Recognit*. 2022(130): 108829.
57. Mabeya H, Khozaim K, Liu T, Orango O, Chumba D, Pisharodi L, Carter J, Cu-Uvin S. Comparison of Conventional Cervical Cytology Versus Visual Inspection with Acetic Acid Among Human Immunodeficiency Virus-infected Women in Western Kenya. *J Low Genit Tract Dis*. 2012; 16(2): 92-97.
58. Mehta V, Vasanth V, Balachandran C. Pap Smear. *Indian J Dermatol Venereol Leprol*. 2009; 75(2): 214.