RESEARCH ARTICLE

VOL. 47 | NO. 1 | JANUARY – APRIL 2026 | PP 1-21

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.47.1.1537

E-LOCATION ID: e1537

Multilayer Perceptron for squamous cell nuclei localization in Pap smear tests using local features

Perceptrón multicapa para la localización de núcleos de células escamosas en pruebas de Papanicolaou usando características locales

Alejandro Reyes Morales[1], *Oscar Susano Dalmau Cedeño*[2],
Teresa Efigenia Alarcón Martínez[1], *Francisco Eduardo Oliva Ibarra*[1]

[1]Centro Universitario de los Valles, Universidad de Guadalajara, Jalisco - México
[2]Centro de Investigación en Matemáticas, CIMAT, A.C, Guanajuato – México

ABSTRACT

Cervical cancer is the fourth most commonly diagnosed cancer among women worldwide and it is a leading cause of death in women. Diagnosis is typically performed manually by analyzing squamous cells in Pap smear tests, a process that is time-consuming and prone to a high rate of errors. The adoption of automated analysis methods has significantly reduced both diagnostic time and human errors. A key step in this automated pipeline is the detection of squamous cell nuclei. In this study, we propose a Multilayer Perceptron (MLP) artificial neural network model that includes local image features to automatically locate squamous cell nuclei. The proposed workflow is intended to detect the presence or absence of nuclei, which aligns with its use in cytology screening. For training we utilized two publicly available datasets: CRIC and SIPaKMeD. Unlike many state-of-the-art approaches, our method learns directly from entire images rather than relying on pre-segmented regions. It achieves accuracies of 0.9526 and 0.9397 on the CRIC and SIPaKMeD datasets, respectively. Our approach offers a simple yet effective solution, delivering performance on par with more complex models reported in the literature.

KEYWORDS: cervical cancer, multilayer perceptron, nuclei, squamous cell

RESUMEN

El cáncer de cuello uterino es el cuarto cáncer más comúnmente diagnosticado entre las mujeres en todo el mundo y sigue siendo una de las principales causas de muerte en las mujeres. El diagnóstico generalmente se realiza manualmente mediante el análisis de células escamosas en pruebas de Papanicolaou, un proceso que requiere mucho tiempo y es propenso a una alta tasa de errores. La adopción de métodos de análisis automatizados ha reducido significativamente tanto el tiempo de diagnóstico como los errores humanos. Un paso clave en este proceso automatizado es la detección de núcleos de células escamosas. En este estudio, se propone un modelo de red neuronal artificial tipo Perceptrón Multicapa (MLP), que utiliza características locales para localizar automáticamente los núcleos de células escamosas. El flujo de trabajo propuesto está diseñado para detectar la presencia o ausencia de núcleos, lo cual está alineado con su uso en el cribado citológico. El modelo fue entrenado con dos conjuntos de datos públicos: CRIC y SIPaKMeD. A diferencia de otros trabajos en el estado del arte, nuestro modelo aprende de la información de imágenes completas alcanzando una precisión de 0.9526 y 0.9397 en los conjuntos de datos CRIC y Sipakmed, respectivamente. Nuestra propuesta ofrece un enfoque sencillo mientras mantiene un rendimiento comparable al de los modelos complejos reportados en la literatura.

PALABRAS CLAVE: cáncer cervicouterino, perceptrón multicapa, núcleos, células escamosas

Corresponding author

TO: **TERESA EFIGENIA ALARCON MARTINEZ**
INSTITUTION: **CENTRO UNIVERSITARIO DE LOS VALLES, UNIVERSIDAD DE GUADALAJARA**
ADDRESS: **CARRETERA GUADALAJARA - AMECA KM. 45.5, AMECA, JALISCO, C.P. 46600, MÉXICO.**
EMAIL: **teresa.alarcon@academicos.udg.mx**

Received:

13 March 2025

Accepted:

13 August 2025

Published:

27 January 2026

INTRODUCTION

Cervical cancer ranks as the fourth most diagnosed cancer in women globally, with around 604,000 new cases and 340,000 deaths reported[1]. Approximately 90% of the 340,000 deaths resulting from this type of cancer occurred in countries with low to moderate income[1][2]. The main cause of this cancer is persistent human papillomavirus infections that cause approximately 95% of cervical cancers. Abnormal cells generally take 15 to 20 years to progress toward cancer, however, in women with a compromised immune system, such as those with untreated HIV, this progression can accelerate, taking only 5 to 10 years to develop into cancer[1][3][4].

The Pap smear test, also known as the Pap test or cervical smear, is the most widely used test to detect cervical cancer. This examination involves gathering cells from the cervix, which are then subjected to various chemical procedures before being sent to clinical laboratories. Subsequently, experts manually identify squamous cells and analyze morphological differences between normal and abnormal cell, by means of a microscope to identify any irregularities that could indicate the presence of cervical cancer or precancerous states[5][6][7]. The diagnosis of cervical cancer through Pap smears focuses on squamous cells, as these are the cells where precancerous and cancerous lesions develop. Pathologists analyze nuclei in these cells because malignant transformations manifest most prominently in the nucleus. Nevertheless, this process is time-consuming and prone to a high incidence of false positives and false negatives due to human errors[5][8]. Additionally, it is highly inefficient, as specialists can only analyze four to five slides per day, leading to excessive workload and visual fatigue[5][7][9]. At the end of the day, diagnoses are significantly delayed, which has negative consequences and critical impact in prognosis and treatment of this pathology[6][10][11]. Therefore, to reduce the negative consequences of manual cell analysis in Pap smear tests, the scientific community has focused on the development of automated systems for segmentation and classification of squamous cells nuclei[12][13][14]. These systems have the capacity to reduce time, minimize the errors, and increase the daily number of analyzed slides[15][16]. By doing so, these systems positively impact the prognosis and treatment of cervical cancer, enabling more timely diagnosis[8][10]. However, the automatic localization of squamous cell nuclei represents a significant challenge due to the complexity and natural variability of Pap smear images. The complexity of the images is attributed to several factors: there are multiple cell types present in the images, leading to detect nuclei that do not belong to squamous cells, furthermore, some artifacts can be introduced to the Pap smear sample during the chemical processes before the manual analysis, and these artifacts might result in a detection of many areas resembling squamous cell nuclei[8][17]. Meanwhile, the scientific community has made numerous contributions in this field, and efforts continue to enhance the performance of different automated diagnostic systems in terms of precision and reliability[8][18]. Nevertheless, the majority of contributions rely on databases of individual cell images, as demonstrated in recent studies conducted by Saguna et al.[19], Maesaroh et al.[20], Chandana et al.[21], Zhang et al.[22]. Such approaches involve a preprocessing step to isolate a single cell from each sensed image, by means of human intervention. This kind of procedure can be more time-consuming than manually analyzing the samples[23][24]. The described challenge and the commented limitation motivate the present research, which handles whole Pap smear images with a variable number of cells, and it presents a practical, automated system tailored to the localization of squamous cell nuclei—a crucial step in Pap smear image analysis. We implement an artificial neural network model for the automatic localization of squamous cell nuclei in whole images from Pap smear tests. The architecture chosen for this task was the Multilayer Perceptron (MLP)[12]. This work is specifically focused on nuclei localization as a preliminary step for subsequent cell analysis and classification for further diagnosis. The experimental setup includes two datasets: CRIC[25] and SIPaKMeD datasets[26].

The remainder of this manuscript is structured as follows: First, we review the state of the art in the classification and segmentation of squamous cell nuclei in Pap smear tests. The Materials and Methods section describes the proposed model, the datasets used, the data generation process, and the evaluation metrics. The Results section presents the experimental findings and their discussion. Finally, we provide the conclusions and outline potential directions for future work.

Related works

The classification and segmentation of areas containing nuclei of squamous cells pose a challenge due to the diversity of cell types present in the images, not limited solely to squamous cells, thus leading to the existence of different types of nuclei. Overlapping or stacked cell regions may be encountered, complicating precise delineation. Additionally, color variations depending on cell depth and factors associated with artifacts contribute to background portions resembling cells[8][18][27].

Some studies in the literature have addressed the issue of classifying nuclei of squamous cells in images derived from Pap smear tests. For example, in the research conducted by Somasundaram et al.[28] the author presents a multi-thresholding algorithm designed to segment the cytoplasm and nucleus regions from the background, the experiments were performed on the Pap-smear Benchmark Data dataset. Morphological operations are applied to rectify the segmented results. Additionally, a support vector machine (SVM) classifier is employed to categorize smear cells as normal or abnormal, based on features extracted from the segmented output.

Wang et al.[8] proposed a Tree Domain structure combined with a screening algorithm to perform nucleus segmentation using CRIC and SIPaKMeD datasets. Initially, the model takes an input image and then proceeds to group regions corresponding to cells. Once the cell regions are grouped, regions of interest are extracted based on depth, upon which the Tree Domain technique is applied to obtain candidate regions for cell nuclei. Finally, the algorithm returns the image with segmented regions belonging to nuclei.

Recently, more novel methods have been utilized for nucleus segmentation. Zhang et al.[22] introduce a new deep neural network (DNN) called the Global Context UNet (GC-UNet), which is specifically designed to effectively address complex settings and provide precise cell segmentation. The essence of GC-UNet lies in DenseNet, which acts as the main component, processing cell images and leveraging existing knowledge. A distinctive pooling module that is sensitive to context, including a gating model, is incorporated to efficiently encode ImageNet dataset pre-trained features, guaranteeing the preservation of crucial features at various levels. Moreover, a decoder with a global context attention mechanism is included to enhance global feature interaction and improve the accuracy of the predicted masks. The experiments were performed on two datasets: the MoNuSeg Dataset and the CoNSep dataset.

The study by Saguna et al.[19] presents an efficient technique called WSAAE-CCD (Water Strider Algorithm with Autoencoder for Cervical Cancer Diagnosis) for the detection of cervical cancer, the authors used the SIPaKMeD Pap smear images dataset. The technique leverages medical image processing and intelligent models, demonstrating time and cost efficiency. The WSAAE-CCD model includes adaptive filtering for noise removal, kernelized fuzzy C-means (KFCM) segmentation to identify affected cervical cells, and a Residual Network (ResNet5) for feature extraction. Finally, an autoencoder-based classification model is optimized using the Water Strider Algorithm (WSA). Experimental results demonstrate that the WSAAE-CCD technique enhances squamous cell nuclei detection performance, thereby contributing to more accurate cervical cancer classification.

The study conducted by Maesaroh et al.[20] focuses on cervical cancer classification using a Convolutional Neural Network (CNN) with transfer learning, based on the Xception architecture. The research allows to classify Pap smear images into two or four categories (normal cells, koilocytes, L-Sil, and H-Sil) using a dataset of 400 images from RepoMedUNM dataset. The methodology involves steps such as data preprocessing, augmentation, hyperparameter tuning, and model evaluation. The CNN model achieved 100% accuracy in the 2-class classification and 85% accuracy in the 4-class classification, demonstrating its effectiveness in detecting and classifying squamous cell nuclei.

In the study conducted in [29], the authors propose ensemble models of convolutional neural networks to extract features and subsequently classify the Pap smear images. In this case, they process entire images, which are then segmented into individual cells for classification. In the method proposed in [30], full images from the CRIC dataset[25] are used. However, a preprocessing step is applied to crop the images into individual cells using the metadata provided by the dataset. Subsequently, the classification is performed at the cellular level.

The work proposed by Chandana et al.[21] explores the use of Deep Learning techniques for automated screening and classification of cervical cancer cells. Utilizing the SIPaKMeD Pap smear image dataset, the proposed method applies the ResNet-50 architecture to classify cervical cells. The results indicate that ResNet-50 achieved the highest accuracy of 91.04%, outperforming both Inception-V3 and MobileNet models. Furthermore, the research integrates Explainable Artificial Intelligence (XAI) techniques, specifically LIME, to enhance the interpretability and transparency of the diagnostic process, thereby improving the reliability of automated squamous cell nuclei and cervical cancer detection.

In this study, we propose an MLP model that utilizes local features for squamous cell nuclei localization in Pap smear images—a critical step in cervical cancer diagnosis. Unlike most state-of-the-art methods, which typically rely on convolutional neural networks (CNNs) with high computational complexity and are trained on small image patches, our approach uses a less complex model and processes entire images. These full images, which may contain a variable number of cells, introduce additional complexity in the task of detecting squamous cell nuclei. We compare our proposed model with both classical and state-of-the-art classification algorithms using two public datasets: CRIC[25] and SiPakMed[26]. Our experimental results demonstrate that the proposed model performs strongly in this task. The proposal constitutes a key step toward the clinical implementation of cytology automation.

MATERIALS AND METHODS

In this section, the details of the proposal are presented. The datasets used are described, and a comprehensive explanation of the two variants of the approach is provided. The rationale behind proposing these two variants is highlighted, along with the advantages of one over the other. Additionally, the inherent complexity of segmenting this type of imagery is discussed. Furthermore, the metrics employed to evaluate the performance of the models are outlined. The training and evaluation phases are also detailed, including the methodologies and procedures followed to ensure robust and reliable results.

Datasets description

The datasets used in this research are enlisted below:

- CRIC dataset comprising 400 images with dimensions of 1376 x 1020 and 150 dpi resolution[25]. CRIC (Center for Recognition and Inspection of Cells) is a collaborative network of researchers dedicated to supplying cell collections to the scientific community. Images were acquired using conventional bright-field

microscopy with a 40× objective and 10× eyepiece, with a Zeiss AxioCam MRc digital camera mounted on a Zeiss AxioImager.Z2 microscope. Certified cytopathologists performed manual identification of squamous cells and classified them in accordance with the Bethesda System nomenclature protocol[36]. The annotations, available in CSV and JSON formats, include the centroid coordinates of each nucleus. This information allows for the extraction of square patches (90–100 pixels per side) centered on the nuclei. According to cytopathological expertise, these regions encompass relevant features of both the nucleus and the adjacent cytoplasm. All images are organized as a collection of Figshare[25] and are licensed under Creative Commons Attribution 4.0 International framework. Figure 1 displays two annotated images taken from the CRIC dataset.

- The SIPaKMeD (Single-cell Images of PAP Smear for Cervical Cancer Detection using KMeans Clustering and Medical Diagnosis) dataset consists of 4,049 high-resolution images (2048 × 1536 pixels) of individually annotated squamous cells, manually cropped from 966 clustered cell images obtained from Pap smear slides, which are also included in the dataset[26]. Image acquisition was carried out using a CCD camera (Infinity 1, Lumenera) mounted on an OLYMPUS BX53F optical microscope. The dataset adopts a non-Bethesda classification for squamous cells. In each image, the cytoplasmic and nuclear regions of the squamous cells are manually outlined by experts. As a result of this procedure the coordinates of each nucleus and cytoplasm are saved in .dat files.

Both datasets are established benchmarks, comprising color images represented in RGB format. All images from the datasets encompass a wide variety of squamous cells, alongside other cell types present in the samples, as well as artifacts. This ensures that experiments are conducted with diverse images, thus avoiding bias in the results with clean or easily classifiable images. The annotated images from the CRIC and SIPaKMeD datasets assists us in generating masks to extract the pixels to be used for the training and testing phases.

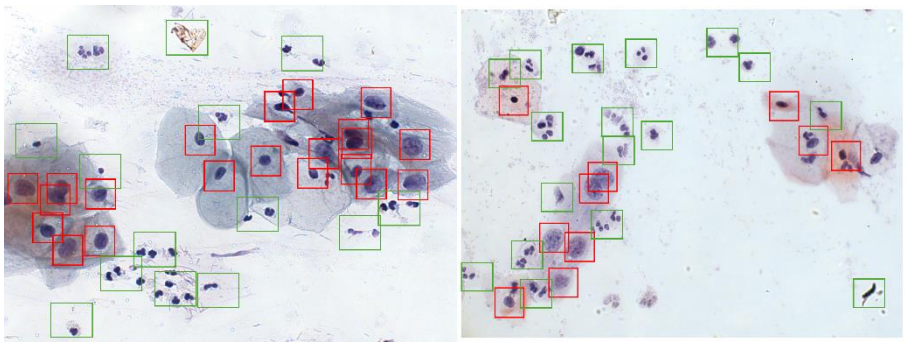

FIGURE 1. **Two images of squamous cells from Pap smear tests, sourced from the CRIC dataset. Red rectangles highlight the annotated squamous cells, while green rectangles mark artifacts and other cell types.**

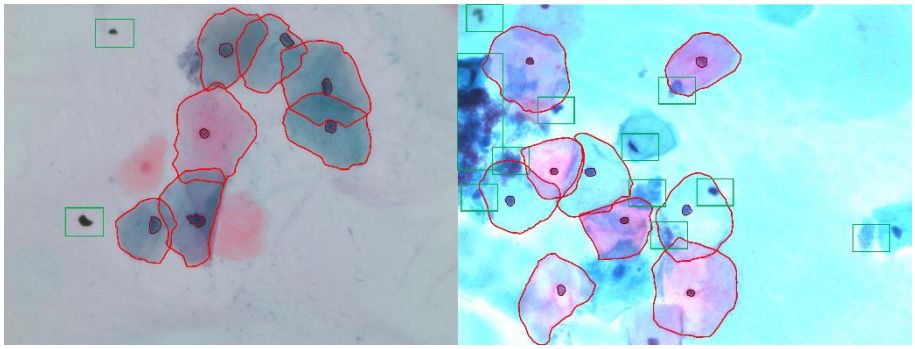

FIGURE 2. **Two images of squamous cells from Pap smear tests, sourced from the SIPaKMeD dataset. Red polygons highlight nuclear and cytoplasmic boundaries using expert-provided coordinates from the dataset. Green rectangles enclose other cell types and artifacts.**

As illustrated in Figures 1 and 2, detecting squamous cell nuclei in whole images poses significant challenges. Figure 1 displays squamous cells enclosed in red rectangles, while Figure 2 shows their nuclei and cytoplasm outlined with red contours. Both images reveal other cell types and artifacts marked by green rectangles, where some artifacts closely resemble squamous nuclei in both morphology and color intensity. Furthermore, nuclei from non-squamous cells exhibit visual similarities that may lead to misclassification.

This complexity suggests that conventional RGB channels alone are insufficient for robust model training. Our approach addresses this by incorporating local pixel-level features specific to nuclear regions. By enriching the feature space with other patterns, the model can better distinguish true squamous nuclei from confounding elements, thereby reducing errors from artifacts or other cell types while maintaining localization sensitivity.

The proposal

The proposal consists of two variants of a Multilayer Perceptron neural network. For the first variant, Figure 3, the MLP only receives information of R, G and B channels for each pixel. We called this model MLP3. The second variant, Figure 4, incorporates additional spatial local characteristics into the model. This model is called MLP33.

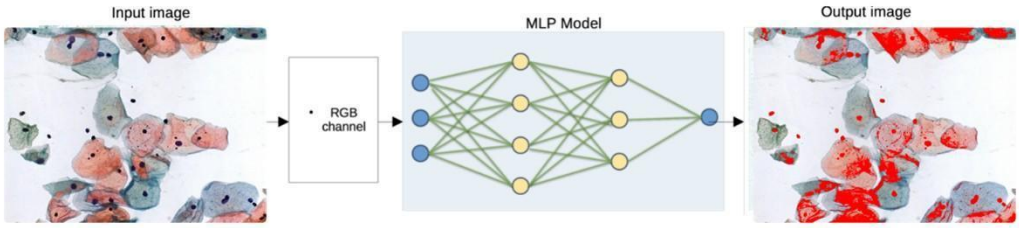

FIGURE 3. **Flowchart illustrating the MLP3 approach for squamous cell nucleus localization using RGB-based features. In the image on the right, red-highlighted regions denote the pixels identified as part of the located nuclei.**

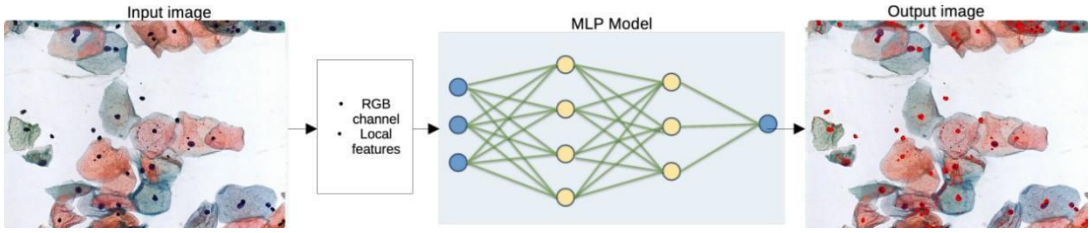

FIGURE 4. **Flowchart illustrating the MLP33 approach for squamous cell nucleus localization using RGB and local features. In the image on the right, red-highlighted regions denote the pixels identified as part of the located nuclei.**

It was observed that the MLP3 adequately classified the pixels in our data set using only their RGB features, see Figure 3. However, some pixels in the target class were misclassified, which was attributed to factors such as the density of cells present in the image, cell overlap, and the presence of artifacts, see Figures 1 and 2. Therefore, leveraging the versatility of the MLP model architecture to introduce complexity, another approach for classifying pixels that represent squamous cell nuclei is proposed, the MLP33. This approach incorporates RGB features along with additional spatial local characteristics into the model, including RGB values of neighboring pixels, as well as the average and standard deviation per RGB channel. The process is described in Figure 4. We consider a neighborhood of 8 pixels.

Data generation for model training and testing

Thirty images were selected from each dataset, and manual segmentation masks were created to extract target pixels (see Figure 5). To extract the features for the MLP3 model, which takes only the three RGB values as input, manually generated masks are used. These masks are created based on data provided by experts for each dataset, which consists of nucleus location coordinates[25][26]. Using this information, the masks are generated to identify nucleus pixels in the images, allowing the extraction of their corresponding RGB values, as illustrated in Figure 6.

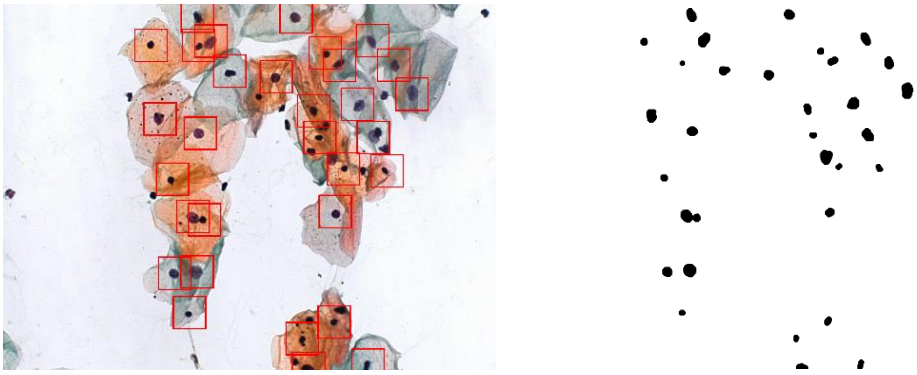

FIGURE 5. **The image on the left represents an example from CRIC dataset. The image on the right corresponds to the generated mask.**

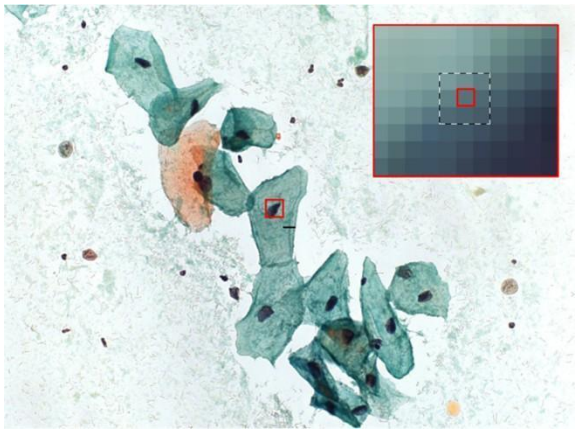

FIGURE 6. **Selecting RGB features from neighboring pixels of the reference pixel (the inner red square).**

To extract the features for the MLP33 model, the RGB values of each pixel labeled as either nucleus or background were considered, along with the RGB values of its eight nearest neighbors, as illustrated in Figure 6. Starting from a labeled pixel (highlighted in red in the image), the RGB values of the surrounding pixels were included to form the training set. In addition, six statistical features were incorporated: the mean and standard deviation for each RGB channel. This results in a total of 33 features used for training. Table 1 provides a summary of the features used in the MLP33 model.

The procedure described above resulted in a total of 1,869,023 and 3,927,565 data for the CRIC and SIPaKMeD sets respectively.

TABLE 1. **Total number of features taken for model training.**

Pixel	Features		
Reference pixel	R1	G1	B1
Neighbor 1	R2	G2	B2
Neighbor 2	R3	G3	B3
Neighbor 3	R4	G4	B4
Neighbor 4	R5	G5	B5
Neighbor 5	R6	G6	B6
Neighbor 6	R7	G7	B7
Neighbor 7	R8	G8	B8
Neighbor 8	R9	G9	B9
Means per channel	Mean R	Mean G	Mean B
Standard deviation per channel	Std R	Std G	Std B
Total per channel	11	11	11
Total features	33		

The MLP model training

The proposed approach for squamous cell nucleus localization is conducted through pixel classification, labeling them as 1 (nucleus class) and 0 (background class). The extracted features for each pixel are used for the training step in correspondence to models MLP3 and MLP33, see Figure 3 and Figure 4.

According to various studies[10][13], using 60–80% of the available data for training is considered appropriate to avoid overfitting. Based on these findings, we split the dataset as follows: 60% for training and 40% for testing. Additionally, increasing the training proportion beyond 60% resulted in significantly longer training times, ranging from 3 to 5 days.

During data preparation for the training phase, we selected samples to ensure balanced class representation, resulting in 934,512 and 934,511 samples for the nucleus and background classes from the CRIC dataset, and 1,963,783 and 1,963,782 samples from the SIPaKMeD dataset.

The MLP3 and MLP33 models were trained for each dataset using the computational resources of the Centro Nacional de Supercómputo of the Instituto Potosino de Investigación Científica y Tecnológica, A.C. Four nodes were employed, each equipped with 192 GB of RAM and 2 processors containing 16 physical cores each. Anaconda with Keras and TensorFlow 2.5 libraries were used for this purpose. Table 2 summarizes the data split for training, validation, and testing, along with the model architecture and its corresponding parameters.

TABLE 2. **MLP configuration and data distribution for the CRIC and SIPaKMeD dataset.**

Parameter	Value
Training Data	60 %
Test Data	40 %
Number of Parameters	12,673
Input Dimension	33
Output Dimension	1
Number of Hidden Layers	2
Nodes per Hidden Layer	[128, 64]
Loss	Binary Cross entropy
Optimizer	Adam
Batch Size	2048
Learning Rate	0.001
Epochs	200

Evaluation metric

To assess the performance of the proposed models, the following metrics were utilized: Accuracy, recall, specificity, and F1 score[31][32][37], see Table 3.

TABLE 3. **Performance metrics for classifier evaluation.**

Metric	Formula	Description
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	Calculates the percentage of accurately classified samples out of the total number of samples.
Recall	$\frac{TP}{TP + FN}$	Measures the proportion of actual positive samples that are correctly identified. Also known as Sensitivity or True Positive Rate.
Specificity	$\frac{TN}{TN + FP}$	Calculates the percentage of true negative samples that are accurately recognized. Also referred to as True Negative Rate.
F1 Score	$2 \times \frac{Precision \times Recall}{Precision + Recall}$ where: $Precision = \frac{TP}{TP + FP}$	Harmonic mean of precision and recall, providing a balance between the two metrics.
IoU	$\frac{Area\ of\ Overlap}{Area\ of\ Union}$	Measures the spatial overlap accuracy between predicted regions and ground truth annotations.

RESULTS AND DISCUSSION

To enhance the classifier's performance, training was conducted by a cross-validation approach and varying the parameters of the layers and nodes within the MLP model to increase precision. It was observed that increasing the number of layers and nodes in the network architecture did not yield a substantial improvement. In fact, in most cases, precision decreased. Consequently, it was decided to maintain the network architecture with two hidden layers, one consisting of 128 nodes and the other of 64 nodes. Additionally, the Adam optimizer with a learning rate of 0.001 was employed. The number of epochs was set to 200, as it was observed that adding more epochs did not lead to further improvements; instead, the model tended to lose generalizability as you can see in Figure 7 and Figure 8. Table 4 shows the numerical results in the test set achieved by MLP3 and MLP33 variants.

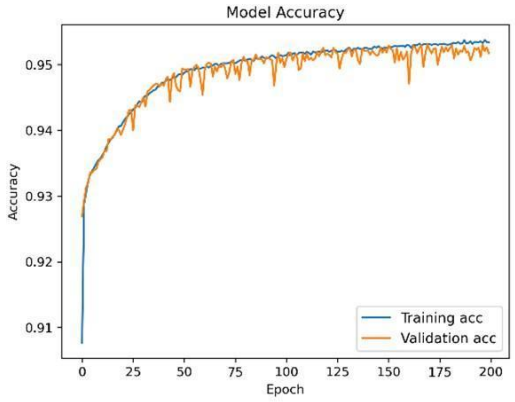

FIGURE 7. **Accuracy plot for the training and validation sets of the CRIC dataset.**

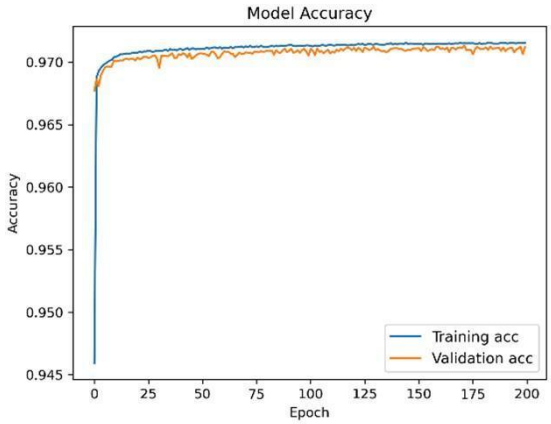

FIGURE 8. **Accuracy plot for the training and validation sets of the SIPaKMeD dataset.**

TABLE 4. **The Accuracy, Recall, Specificity and F1 Scores values for two variants of the proposed model in the test set: MLP3 and MLP33.**

Dataset	Model	Accuracy	Recall	Specificity	F1 Score
CRIC	MLP33	**0.9526**	**0.9695**	**0.9372**	**0.9530**
	MLP3	0.9255	0.9271	0.9084	0.9271
SIPaKMeD	MLP33	**0.9397**	**0.9614**	**0.9213**	**0.9410**
	MLP3	0.9255	0.9345	0.9082	0.9279

Based on the experimental results shown in Table 4, we selected the MLP33 model, described in Section "The Proposal", which incorporates not only RGB features but also local characteristics, along with channel-wise averages and standard deviations. It is worth noting that for the two datasets considered, the MLP33 model consistently yields higher evaluation metrics.

Cross-validation performance during training

To ensure fair benchmarking against our proposed model, we additionally trained classical machine learning models using the same features. All models were implemented in Python with Keras and Scikit-learn, following identical data preprocessing steps and the same 60/40 train-test split. We implemented the 5-fold cross-validation in order to analyze the overfitting of all studied models.

The classical classifiers included in this study, were:

- Support Vector Machine (SVM) with Radial Basis Kernel.
- K-Nearest Neighbors (KNN) with the number of neighbors set to 4.
- Decision Trees using the default configuration provided by Keras.
- Random Forest with 50 trees in the ensemble.
- Bayesian Gaussian Mixture Model (GMM) with 4 Gaussian components in the mixture.

Tables 5 and 6 present the cross-validation results, for the CRIC and SIPaKMeD datasets respectively.

TABLE 5. **Cross-validation performance on the CRIC dataset: mean metrics ± standard deviation**

Model	Mean accuracy on training	Mean Accuracy on validation
MLP33	0.9504 ± 0.0015	0.9498 ± 0.0014
SVM	0.9318 ± 0.0001	0.9317 ± 0.0005
K Nearest Neighbor	0.9589 ± 0.0001	0.9460 ± 0.0006
Decision Trees	0.9247 ± 0.0002	0.9243 ± 0.0003
Random Forest	0.9998 ± 0.0000	0.9469 ± 0.0007
Bayesian Gaussian mixture	0.8874 ± 0.0001	0.8872 ± 0.0005

TABLE 6. **Cross-validation performance on the SIPaKMeD dataset on train set: mean metrics ± standard deviation**

Model	Mean accuracy on training	Mean Accuracy on validation
MLP33	0.9316 ± 0.0007	0.9315 ± 0.0009
SVM	0.9312 ± 0.0005	0.9310 ± 0.0004
K Nearest Neighbor	0.9388 ± 0.0002	0.9388 ± 0.0002
Decision Trees	0.8937 ± 0.0002	0.8934 ± 0.0002
Random Forest	0.9998 ± 0.0000	0.9382 ± 0.0003
Bayesian Gaussian mixture	0.8908 ± 0.0039	0.8909 ± 0.0041

As shown in Tables 5 and 6 the training results for the CRIC and SIPaKMeD datasets respectively demonstrate consistent performance across most classifiers. The low standard deviations across cross-validation folds reflect consistent metric performance, providing evidence of model stability. For both the CRIC and SIPaKMeD datasets, the average performance metrics are satisfactory, indicating good generalization and no evidence of overfitting in the

training set.

Final evaluation on test set

Below we compare the MLP33 with other classical methods and state-of-the-art models from the reviewed literature. No preprocessing was performed prior to training to ensure a fair comparison. Table 7 and Table 8 present the comparison considering the CRIC and SIPaKMeD datasets, on the test set respectively.

The state-of-the-art results used for comparison were taken directly from the reported findings in the original research papers, which did not require any additional implementation to replicate these results.

TABLE 7. Performance comparison of the MLP33 model with classical and state-of-the-art classifiers on the CRIC dataset.

Model	Accuracy	Recall	Specificity	F1 score
MLP33	0.9526	0.9695	0.9372	0.9530
SVM	0.9262	0.9262	0.9033	0.9279
K Nearest Neighbor	0.9464	0.9576	0.9351	0.9471
Decision Trees	0.9158	0.913	0.9185	0.9155
Random Forest	0.9472	0.9698	0.9246	0.9483
Bayesian Gaussian mixture	0.9029	0.9244	0.8814	0.9050
DL-Ensemble, Daniz *et. al.*[29]	0.9600	0.9600	0.9600	0.9600
Hierarchical Feature-Based, Daniz *et. al.*[30]	0.9450	0.8348	0.9670	0.8351
Shape-based Features,Terra *et. al.*[33]	0.9420	0.9510	0.9170	0.9600

TABLE 8. Performance comparison of the MLP33 model with classical and state-of-the-art classifiers on the SIPaKMeD dataset.

Model	Accuracy	Recall	Specificity	F1 score
MLP33	0.9397	0.9614	0.9213	0.9410
SVM	0.9395	0.9608	0.9321	0.9355
K Nearest Neighbor	0.9392	0.9565	0.9219	0.9401
Decision Trees	0.9003	0.8973	0.9032	0.8999
Random Forest	0.9387	0.9589	0.9185	0.9399
Bayesian Gaussian mixture	0.8215	0.9584	0.6845	0.843
CNN+PCA, GV *et. al.*[34]	0.9637	N/A	0.9983	0.9963

As shown in Tables 7 and 8, the model consistently achieves high recall values (0.9695 for CRIC and 0.9614 for SIPaKMeD), indicating its effectiveness in identifying the majority of nucleus pixels across different datasets. Although false positives are present, the F1 scores remain high (0.9530 for CRIC and 0.9410 for SIPaKMeD), reflecting a favorable balance between sensitivity and precision. In the context of cytology screening, this trade-off is acceptable—and even desirable—as such applications typically prioritize sensitivity over specificity to minimize the risk of missing clinically relevant cells.

Although full images are not analyzed in [29], this study is included to highlight the robustness of the proposed model. It is important to note that, while the reviewed state-of-the-art studies report higher accuracy, they rely on models that use individual cells, even though these cells are derived from whole images. This distinction directly impacts the complexity of the problem. Our proposed method, which processes the entire image without the preprocessing and cropping steps used in [30], demonstrates higher performance. This confirms the relevance of

our approach for processing full images.

As observed in [34], their model achieves higher accuracy than ours; however, their approach does not process entire images. Instead, they randomly crop an unspecified area from the original image, which is then resized to 224x224 for processing with the ResNet-34 network[35]. This method excludes a significant portion of the original image, leaving many cells out of the analysis. Therefore, it does not operate on fully intact images, but rather on "semi-complete" ones. For the SIPaKMeD dataset, only one state-of-the-art result is mentioned[34], as an exhaustive search revealed no additional literature utilizing full-slide images. Most studies focus solely on single-cell images rather than on entire images.

In short, after analyzing the results and comparing them with state-of-the-art findings, it can be determined that our proposed model, performs comparably to the results reported in the literature. Notably, these studies often employ more complex and computationally expensive models. Our approach demonstrates that it is possible to achieve competitive performance without resorting to the increased complexity of those models.

To validate the robustness of our MLP33 architecture, we conducted cross-dataset testing. The model trained on the CRIC dataset achieved 85.91 % accuracy, recall, and specificity, along with an F1-score of 86.36 % on the SIPaKMeD test set (see Table 9), indicating a moderate level of domain transfer. This represents an 8.9% drop in F1-score compared to its original performance on the CRIC test set (95.26%). In contrast, the model trained on the SIPaKMeD dataset exhibited stronger generalization to the CRIC test set, attaining 90.69% accuracy, 93.53% recall, 87.69% specificity, and an F1-score of 90.95% (see Table 10). These findings suggest that the broader morphological variability within SIPaKMeD enhances the model's ability to learn transferable features, thereby improving cross-domain performance.

TABLE 9. **Cross-dataset performance of CRIC-trained MLP33**

Test set	Accuracy	Recall	Specificity	F1 Score
CRIC (native)	0.9526	0.9695	0.9372	0.9530
SIPaKMeD	0.8591	0.8591	0.8591	0.8636

TABLE 10. **Cross-dataset performance of SIPaKMeD-trained MLP33**

Test set	Accuracy	Recall	Specificity	F1 Score
SIPaKMeD (native)	0.9397	0.9614	0.9213	0.9410
CRIC	0.9069	0.9353	0.8786	0.9095

Figure 9 depicts the classification results of the MLP33 model. Pixels classified as nucleus are highlighted in red. Comparing the classifier's results shown in Figure 9(e) and Figure 9(f) with the ground truth masks in Figure 9 (c) and Figure 9 (d), we can observe certain areas where pixels are incorrectly classified as nucleus (some of these misclassified areas are enclosed in black squares). These misclassified pixels represent background areas mistakenly identified as nucleus, however this does not substantially affect the determination of a nucleus region, since the errors occur at pixel level.

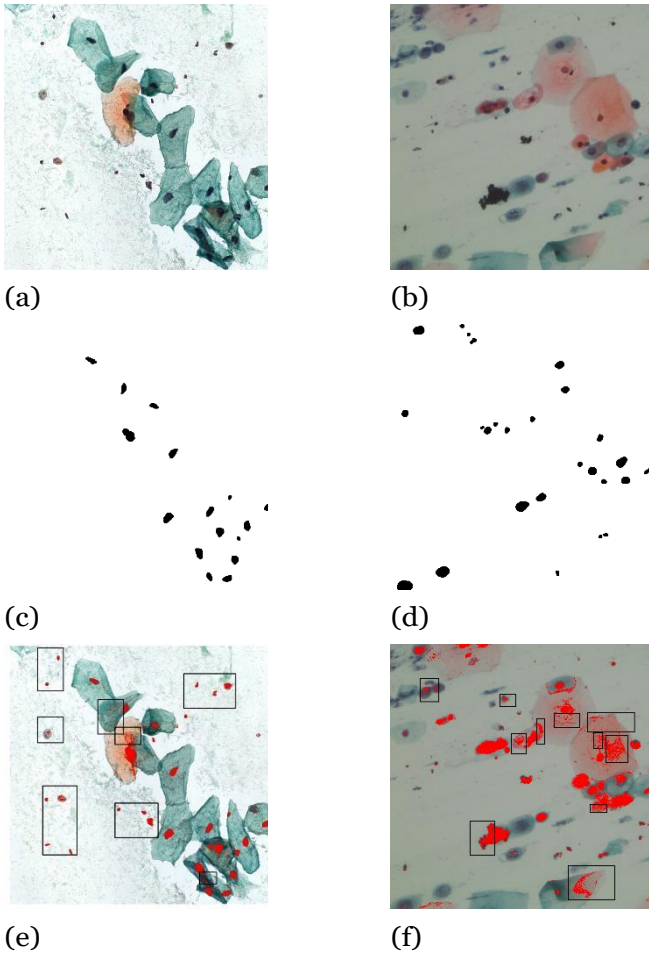

FIGURE 9. **Graphical representation of the proposal: panels (a) and (b) contain images examples from CRIC, and SIPaKMeD datasets, respectively. Images in the second row, (c) and (d) correspond to the ground truth of images in the first row. Images in panels (e) and (f) represent the overlaid classification results onto the images (a) and (b).**

It was observed that the proposed model performed well despite the variety and complexity of the data. The model achieved better performance on the CRIC dataset. It had more difficulty classifying nucleus pixels in the SIPaKMeD dataset due to the larger variety of images and characteristics, see Figure 9(f) and Figure 9(e).

Figure 10 presents additional visual results of the MLP33 model on two images: one from the CRIC dataset and another from the SIPaKMeD dataset.

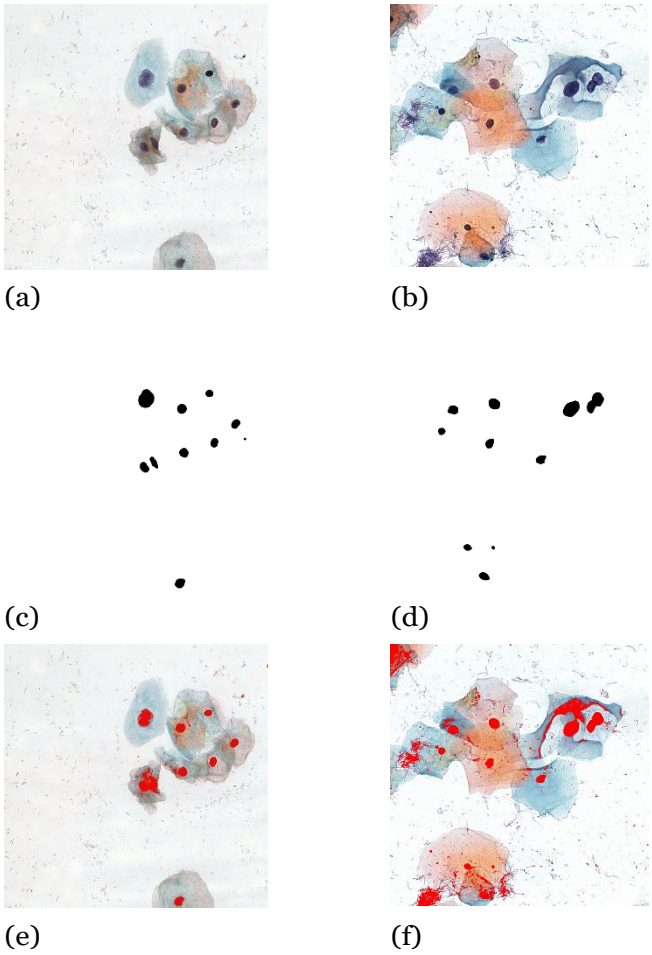

FIGURE 10. **Graphical representation of the proposal: panels (a) and (b) contain images examples from CRIC, and SIPaKMeD datasets, respectively. Images in the second row, (c) and (d) correspond to the ground truth of images in the first row. Images in panels (e) and (f) represent the overlaid classification results onto the images (a) and (b).**

To further evaluate the performance of our proposal, we employ the Intersection over Union (IoU) metric. Thirty images from the CRIC and SIPaKMeD datasets were analyzed using the MLP33 model. The Intersection over Union (IoU) metric was computed, resulting in average IoU scores of 0.5214 for CRIC and 0.4622 for SIPaKMeD. These results are presented in Table 11.

TABLE 11. **Assessment of MLP33 performance using IoU on the CRIC and SIPaKMeD datasets.**

Metric	CRIC Dataset	SIPaKMeD Dataset
Min IoU	0.3415	0.3187
Max IoU	0.8909	0.6142
Mean IoU	0.5214	0.4622
Std Dev	±0.1436	±0.0896

The results indicate that our approach does not precisely capture the morphology of nuclei, as it performs semantic segmentation rather than object-level segmentation. This distinction is clarified below:

- The MLP model was designed for nuclei localization rather than fine-grained segmentation. Its primary goal is to identify regions containing nuclei (to enable subsequent bounding box extraction; see the Results section), without aiming for pixel-perfect boundaries. This design rationale supports our use of pixel-wise

classification metrics—Accuracy, Recall, Specificity, and F1-score—instead of region-based metrics such as Intersection over Union (IoU).

- Classification errors made by the MLP can introduce segmentation artifacts that are penalized by the IoU metric. Nevertheless, the model effectively enables nuclei detection, even when boundary delineation is imperfect. While IoU emphasizes contour precision, our workflow focuses on determining the presence or absence of nuclei, which aligns with the intended purpose of our model.

Figure 11 provides a visual illustration of these points.

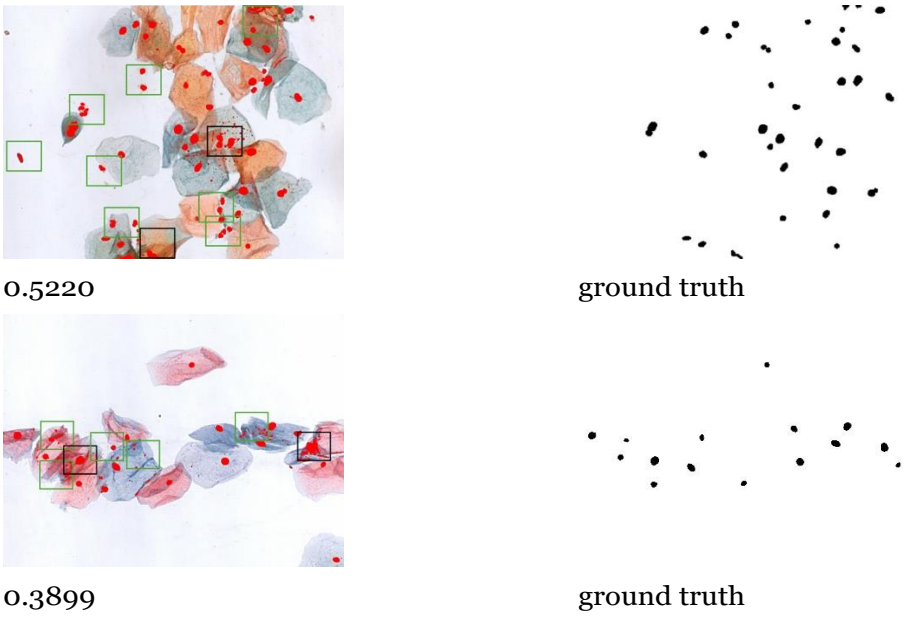

FIGURE 11. **Left panel: MLP33 model predictions evaluated using the IoU metric. Misclassified artifacts and regions are marked with green bounding boxes, and segmentation errors are indicated with black bounding boxes. Right panel: corresponding ground truth masks.**

The examples in the Figure 11 illustrate that, while the full extent of each nucleus may not be captured, the model successfully identifies their presence or absence. Red areas denote the classifier's results. Green bounding boxes mark misclassified artifacts and regions, whereas black bounding boxes indicate nuclei with segmentation errors either over-segmentation (excessive area) or under-segmentation (incomplete coverage).

The proposed model demonstrates a more efficient classification of nucleus pixels compared to the other models. Figure 12 illustrates that the proposed model outperforms others in classifying nucleus pixels. The prediction in Figure 12(c) shows the best alignment with the ground truth mask in Figure 12(b). Although some small background areas are misclassified, the segmentation of the nuclei is more consistent with the ground truth mask. This becomes even more evident when comparing the ground truth mask with the results of the other classifiers. For example, a preliminary inspection reveals that both the Support Vector Machine and the Bayesian classifier incorrectly labeled a substantial portion of the background as nucleus regions. Furthermore, upon closer inspection of the other models, small artifacts found in the background are classified as part of the nuclei, demonstrating the generalization capability of our proposed model in pixel classification tasks, considering local features.

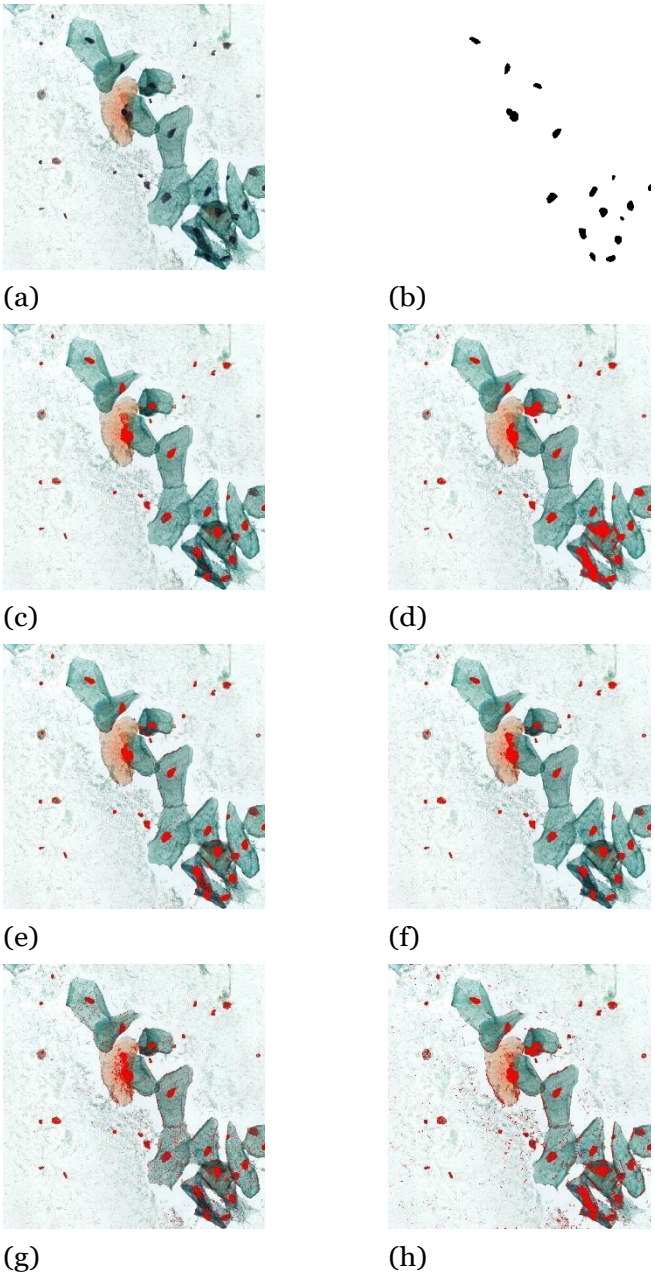

FIGURE 12. **A graphical comparison of our model with other classification models. (a) Image example from CRIC dataset, (b) the ground truth image, classification results of (c) our model, (d) SVM, (e) K-nearest neighbors, (f) Random Forest, (g) Decision tree, and (h) Bayes classifier.**

After detecting the nuclei, the subsequent stage of our research involves postprocessing and connected component analysis. In alignment with practices used by cytopathologists in the CRIC dataset, we intend to extract square patches (90–100 pixels per side) centered on the centroid of each detected nucleus. This will enable the model to capture both nuclear and cytoplasmic features, supporting more comprehensive automated analysis of Pap smear images.

CONCLUSIONS AND FUTURE WORK

In this work was proposed an Artificial Neural Network model based on the Multilayer Perceptron, for the classification of pixels belonging to the nucleus of squamous cells. RGB characteristics were considered, and additionally, local characteristics were added, which include: RGB components of their 8 nearest neighbors, channel-wise averages and standard deviations. Furthermore, two public datasets were considered. The proposed model achieved an accuracy, recall, specificity, and F1 score of 0.9526, 0.9695, 0.9372, and 0.9530, respectively, on the CRIC dataset; and 0.9397, 0.9614, 0.9213, and 0.9410 on the SIPaKMeD dataset.

Our approach enables automatic pixel-wise localization of squamous cell nuclei directly on full Pap smear images, without the need for manual cell isolation or cropping. The proposed model advances the development of clinically viable automated screening solutions.

It was confirmed that classical supervised learning models remain useful in complex tasks like the one addressed by this proposed model. We observed that the performance metrics surpass those of the reviewed literature using more complex and resource-intensive models. Therefore, it is important to continue considering these classical models when performing complex classification tasks, such as segmenting nuclei of squamous cells to automate cervical cancer diagnoses.

As future work, alternative approaches are proposed to enhance the performance in the classification of squamous cell nuclei. Additionally, possible improvements can be achieved by postprocessing, connected component analysis and bounding box extraction on the obtained results. On the other hand, the use of convolutional neural networks, particularly U-Net architectures, it is suggested to perform nucleus segmentation. Subsequently, other models will be applied to characterize whether cells are normal or abnormal based on the segmentation results, aiming to automatically detect malignant cells within Pap smear medical tests.

ACKNOWLEDGMENTS

We would like to express our gratitude to the National Council of Humanities, Sciences and Technologies for the scholarship awarded to me for the completion of my doctoral studies and, consequently, for the completion of this project. Also extend our gratitude to the National Supercomputing Center at the Potosino Institute of Scientific and Technological Research A.C, for providing us with the opportunity to utilize their high-performance computing resources for the completion of this project.

REFERENCES

[1] 073 Regional Committee for the Western Pacific, "Cervical cancer." WHO Regional Office for the Western Pacific, p. Provisional agenda item 10, 2022.

[2] 72nd session Regional Committee for Europe, "Seventy-second Regional Committee for Europe: Tel Aviv, 12– 14 September 2022: roadmap to accelerate the elimination of cervical cancer as a public health problem in the WHO European Region 2022–2030." World Health Organization. Regional Office for Europe, p. 8 p., 2022.

[3] G. Bogani et al., "Duration of human papillomavirus persistence and its relationship with recurrent cervical dysplasia," Eur. J. Cancer Prev., vol. 32, no. 6, pp. 525–532, 2023.

[4] J. Lu, S. Han, J. Na, Y. Li, J. Wang, and X. Wang, "The correlation between multiple HPV infections and the occurrence, development, and prognosis of cervical cancer," Front. Microbiol., vol. 14, p. 1220522, 2023.

[5] L. A. Liang et al., "Cervical cancer screening: comparison of conventional Pap smear test, liquid-based cytology, and human papillomavirus testing as stand-alone or cotesting strategies," Cancer Epidemiol. Biomarkers Prev., vol. 30, no. 3, pp. 474–484, 2021.

[6] M. T. Corkum et al., "When pap testing fails to prevent cervix cancer: a qualitative study of the experience of screened women under 50 with advanced cervix cancer in Canada," Cureus, vol. 11, no. 1, 2019.

[7] P. Dilip Nandanwar, V. M. Wadhai, A. S. Chanchlani, and V. M. Thakare, "Analysis of Pixel Intensity Variation by Performing Morphological Operations for Image Segmentation On Cervical Cancer Pap Smear Image," in 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), Nov. 2021, pp. 1–6. doi: https://doi.org/10.1109/ICCICA52458.2021.9697185

[8] N. A. Alias, W. A. Mustafa, M. A. Jamlos, S. Ismail, H. Alquran, and M. N. K. H. Rohani, "Pap Smear Image Analysis Based on Nucleus Segmentation and Deep Learning–A Recent Review," J. Adv. Res. Appl. Sci. Eng. Technol., vol. 29, no. 3, pp. 37–47, 2023.

[9] S. Haridas and J. T, "Chaos and Exponential Scale based Butterfly Optimization Technique for Feature Extraction and Selection in Pap Smear Images," in 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), Mar. 2023, pp. 740–744. Accessed: Feb. 04, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10112370

[10] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in Pap smears," Comput. Methods Programs Biomed., vol. 113, no. 2, pp. 539–556, 2014, doi: https://doi.org/10.1016/j.cmpb.2013.12.012

[11] S. Dasgupta, "The Efficiency of Cervical Pap and Comparison of Conventional Pap Smear and Liquid-Based Cytology: A Review," Cureus, vol. 15, no. 11, 2023.

[12] M. Sangworasil et al., "Automated Screening of Cervical Cancer Cell Images," in 2018 11th Biomedical Engineering International Conference (BMEiCON), Nov. 2018, pp. 1–4. doi: https://doi.org/10.1109/BMEiCON.2018.8609958

[13] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Segmentation and Classification Techniques for Pap Smear Images in Detecting Cervical Cancer: A Systematic Review," IEEE Access, vol. 12, pp. 118195–118213, 2024, DOI: https://doi.org/10.1109/ACCESS.2024.3447887

[14] W. William, A. H. Basaza-Ejiri, J. Obungoloch, and A. Ware, "A Review of Applications of Image Analysis and Machine Learning Techniques in Automated Diagnosis and Classification of Cervical Cancer from Pap-smear Images," in 2018 IST-Africa Week Conference (IST-Africa), May 2018, p. Page 1 of 11-Page 11 of 11. Accessed: Feb. 04, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/8417373

[15] N. Sompawong et al., "Automated Pap Smear Cervical Cancer Screening Using Deep Learning," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2019, pp. 7044–7048. doi: https://doi.org/10.1109/EMBC.2019.8856369

[16] J. S and D. F. X. Christopher, "Battle Royale Optimization Algorithm for Automatic Threshold Selection in Cervical Pap Smear Images," in 2023 International Conference on IoT, Communication and Automation Technology (ICICAT), Jun. 2023, pp. 1–5. doi: https://doi.org/10.1109/ICICAT57735.2023.10263770

[17] Z. Lu et al., "Evaluation of Three Algorithms for the Segmentation of Overlapping Cervical Cells," IEEE J. Biomed. Health Inform., vol. 21, no. 2, pp. 441–450, Mar. 2017, doi: https://doi.org/10.1109/JBHI.2016.2519686

[18] N. Youneszade, M. Marjani, and C. P. Pei, "Deep Learning in Cervical Cancer Diagnosis: Architecture, Opportunities, and Open Research Challenges," IEEE Access, vol. 11, pp. 6133–6149, 2023, doi: https://doi.org/10.1109/ACCESS.2023.3235833

[19] C. Suguna and S. P. Balamurugan, "An Efficient Water Strider Algorithm with Auto Encoder for Cervical Cancer Diagnosis using Pap Smear Images," in 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), Jan. 2022, pp. 1463–1468. doi: https://doi.org/10.1109/ICSSIT53264.2022.9716358

[20] I. S. Maesaroh, K. N. Syaja'Ah, Y. S. Perkasa, S. Hadianti, D. Riana, and R. R. Nurmalasari, "Cervical Cancer Classification From Pap Smear Using XCeption Model," in 2024 10th International Conference on Wireless and Telematics (ICWT), Jul. 2024, pp. 1–5. doi: https://doi.org/10.1109/ICWT62080.2024.10674721

[21] B. S. Chandana, C. Kommana, G. S. Madhav, P. Basa Pati, T. Singh, and A. K, "Explainable Screening and Classification of Cervical Cancer Cells with Enhanced ResNet-50 and LIME," in 2024 3rd International Conference for Innovation in Technology (INOCON), Mar. 2024, pp. 1–7. doi: https://doi.org/10.1109/INOCON60754.2024.10512322

[22] E. Zhang et al., "Cervical cell nuclei segmentation based on GC-UNet," Heliyon, vol. 9, no. 7, 2023.

[23] H. Kaur, R. Sharma, and L. Kaur, "Automated Cervical Cancer Image Analysis using Deep Learning Techniques from Pap-Smear Images: A Literature Review," in 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Sep. 2021, pp. 1–7. doi: https://doi.org/10.1109/ICRITO51393.2021.9596102

[24] D. Jia, Z. Li, and C. Zhang, "A Parametric Optimization Oriented, AFSA Based Random Forest Algorithm: Application to the Detection of Cervical Epithelial Cells," IEEE Access, vol. 8, pp. 64891–64905, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2984657

[25] M. T. Rezende et al., "Cric searchable image database as a public platform for conventional pap smear cytology data," Sci. Data, vol. 8, no. 1, pp. 1–8, 2021.

[26] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, "SIPaKMeD: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images," in 2018 25th IEEE International Conference on Image Processing (ICIP), 2018, pp. 3144–3148. doi: https://doi.org/10.1109/ICIP.2018.8451588

[27] Z. Meng, Z. Zhao, F. Su, and W. Wang, "Adaptive Elastic Loss Based on Progressive Inter-Class Association for Cervical Histology Image Segmentation," in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2020, pp. 976–980. doi: https://doi.org/10.1109/ICASSP40776.2020.9053232

[28] D. Somasundaram, S. Gnanasaravanan, and N. Madian, "Automatic segmentation of nuclei from pap smear cell images: A step toward cervical cancer screening," Int. J. Imaging Syst. Technol., vol. 30, no. 4, pp. 1209–1219, 2020, doi: https://doi.org/10.1002/ima.22444

[29] D. N. Diniz et al., "A Deep Learning Ensemble Method to Assist Cytopathologists in Pap Test Image Classification," J. Imaging, vol. 7, no. 7, Art. no. 7, Jul. 2021, doi: https://doi.org/10.3390/jimaging7070111

[30] D. N. Diniz et al., "A Hierarchical Feature-Based Methodology to Perform Cervical Cancer Classification," Appl. Sci., vol. 11, no. 9, Art. no. 9, Jan. 2021, doi: https://doi.org/10.3390/app11094091

[31] "Introduction to Machine Learning," MIT Press. Accessed: Sep. 25, 2024. [Online]. Available: https://mitpress.mit.edu/9780262043793/introduction-to-machine-learning/

[32] C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.

[33] D. Terra, A. Lisboa, M. Rezende, C. Carneiro, and A. Campos, "Shape-based Features Investigation for Preneoplastic Lesions on Cervical Cancer Diagnosis," Feb. 2023. doi: https://doi.org/10.5220/0011900800003417

[34] K. K. GV and G. M. Reddy, "Automatic Classification of Whole Slide Pap Smear Images Using CNN With PCA Based Feature Interpretation," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun. 2019, pp. 1074–1079. doi: https://doi.org/10.1109/CVPRW.2019.00140

[35] S. Targ, D. Almeida, and K. Lyman, "Resnet in Resnet: Generalizing Residual Architectures," Mar. 25, 2016, arXiv: arXiv:1603.08029. doi: https://doi.org/10.48550/arXiv.1603.08029

[36] National Cancer Institute (NCI), "The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes", 3rd ed. New York: Springer, 2015.

[37] J. Long et al., "Fully Convolutional Networks for Semantic Segmentation," Proc. IEEE CVPR, pp. 3431–3440, 2015.