

Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization

Editor
João Manuel R.S. Tavares

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tciv20>

Segmentation of cervical nuclei using convolutional neural network for conventional cytology

Júlia Beatriz A. Teixeira, Mariana T. Rezende, Débora N. Diniz, Cláudia M. Carneiro, Eduardo J. da S. Luz, Marcone J. F. Souza, Daniela M. Ushizima, Fátima N. S. de Medeiros & Andrea G. Campos Bianchi

To cite this article: Júlia Beatriz A. Teixeira, Mariana T. Rezende, Débora N. Diniz, Cláudia M. Carneiro, Eduardo J. da S. Luz, Marcone J. F. Souza, Daniela M. Ushizima, Fátima N. S. de Medeiros & Andrea G. Campos Bianchi (2023) Segmentation of cervical nuclei using convolutional neural network for conventional cytology, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11:5, 1876-1888, DOI: [10.1080/21681163.2023.2197086](https://doi.org/10.1080/21681163.2023.2197086)

To link to this article: <https://doi.org/10.1080/21681163.2023.2197086>



Published online: 05 Apr 2023.



Submit your article to this journal



Article views: 63



View related articles



View Crossmark data

Segmentation of cervical nuclei using convolutional neural network for conventional cytology

Júlia Beatriz A. Teixeira^a, Mariana T. Rezende^b, Débora N. Diniz^a, Cláudia M. Carneiro^b, Eduardo J. da S. Luz^a, Marcone J. F. Souza^a, Daniela M. Ushizima^c, Fátima N. S. de Medeiros^d and Andrea G. Campos Bianchi^a

^aDepartamento de Computação, Universidade Federal de Ouro Preto, Ouro Preto-MG, Brazil; ^bDepartamento de Análises Clínicas, Universidade Federal de Ouro Preto, Ouro Preto-MG, Brazil; ^cDepartment of Energy, University of California - Berkeley, Berkeley CA, USA; ^dDepartamento de Engenharia de Teleinformática, Universidade Federal do Ceará, Fortaleza, CE, Brazil

ABSTRACT

Although implementing the Pap smear has drastically reduced the mortality rates from cervical cancer, false positives and negatives are related to the quality of the analysis and the cytopathologist experience. An alternative is the insertion of digital cytology in the quality monitoring to assist the screening. However, conventional cytology is still a major challenge, as it presents a lot of cellular overlap and several epithelial structures that make it difficult to implement computational methodologies. This article compares the performance of U-net and SegNet neural networks for nuclei segmentation in cervical images. Experiments were performed with different activation functions, batch sizes, and datasets, ISBI (synthetic images from liquid cytology) and CRIC Cervix-Seg (conventional cytology real images). The models achieved a Dice coefficient of 0.9783 for ISBI2015 and 0.9429 for CRIC Cervix-Seg. These results suggest a methodology capable of segmenting real images of cervical nuclei with quality, even in situations of overlap and artefacts, advancing efforts towards the automation of tasks as part of the cytopathological analysis in the laboratory work routine.

ARTICLE HISTORY

Received 17 October 2022
Accepted 20 March 2023

KEYWORDS

Cervical nuclei;
segmentation; convolutional
neural network; pap smear

1. Introduction

Since the implementation of the Pap smear in the 1950s to screen for cervical cell abnormalities, mortality, and incidence rates have dramatically decreased (Liang et al. 2021). Nevertheless, cervical cancer remains the fourth most frequent cause of cancer death in women worldwide (WHO 2022). It may be prevented through early diagnosis using screening and treatment of precancerous lesions (Chisale Mabotja et al. 2021). Screening detects cancer at an early stage where treatment has a high potential for cure.

In low-and middle-income countries, tests rely on visual inspection. After collection, the material is sent to the laboratory for visual inspection. Under the optical microscope, professionals interpret the observed image, which is done manually through the recognition and identification of cellular structures, following internationally adopted protocols, called the Bethesda System (TBS), for the assessment and description of cellular lesions (Nayar and Wilbur 2017). TBS stands out as a standardised report model in cervicovaginal cytology (Pangarkar 2022). In addition to its reproducibility, it reflects the most current understanding of cervical cancer. The most important feature is its clinical relevance. Each classification category has clear clinical implications based on solid evidence and worldwide consensus.

Despite the constant updating of professionals and the assessment protocols required by the scientific community, conventional cytology still has limitations due to the number of cells, inflammatory agents, and overlaps. Meantime, liquid-

based cytology images are more homogeneous and easier to analyse. The visual interpretation of a conventional cytology slide is time-consuming, involves interpretative evaluations, and requires much technical knowledge from the professional. Therefore to improve the quality of the results and support the professional service, it is possible to include methodologies to augment and rank the different lesions for later cytopathologist evaluation. In computer-assisted cervical cancer screening, nuclei detection is introduced to facilitate image analysis by extracting meaningful information.

Several authors argue that the cell nucleus contains the vast majority of information for determining the lesion (Nayar and Wilbur 2017; Zou et al. 2020; Diniz et al. 2021) since they present significant morphological and textural variations, responsible for providing essential information to the diagnosis. Therefore, our hypothesis relies on the importance of cervical nuclei segmentation to improve assertiveness in diagnosis since its delimitation can be an important step for cell lesion characterisation.

This article compares the performance of two deep learning convolutional neural network architectures (U-net and Segnet) to segment cell nuclei in real images of cervical cells obtained from conventional Pap smears. The architectures performance was evaluated using ISBI2015 and CRIC Cervix-Seg databases. We also highlight the proposal of CRIC Cervix-Seg, not used in previous publications, a segmentation database based on conventional cytology with images of cervical cells from CRIC Cervix (Rezende et al. 2021). Once the cervical nuclei have



been correctly identified, they can be used in intelligent digital teleytology systems for classification and decision support for professionals during cervical cancer diagnoses.

The main contributions of this work can be summarised as follows: (i) a proposal of a convolutional neural network for automatic segmentation of cell nuclei in real Pap smear images; (ii) the use of multi-cells images of conventional cytology instead of isolated ones; (iii) the publication of a representative segmentation dataset for cell nuclei (3,200 segmented cell nuclei obtained from 400 images); and, finally, (iv) Dice state-of-the-art results for the ISBI2015 and CRIC Cervix-Seg segmentation collection.

The outline of the paper is as follows. [Section 2](#) presents the recent works and databases used on the problem under study. [Section 3](#) exhibits the materials and methods considered. [Section 3.4.2](#) displays the computational experiments and their results and discussions. Finally, [Section 3.4.2](#) presents the conclusions of this work.

2. Related works

This section presents some related works for cervical cell segmentation. It begins with a summary of the main databases available for cell segmentation, followed by a description of the deep convolutional methods considered for comparing and investigating the theoretical foundation. Finally, a discussion of the approaches proposed by the authors was presented.

2.1. Segmentation databases

Several authors have addressed cervical cell segmentation involving databases containing isolated, synthetic cells or liquid-based cytology. According to [Jiang et al. \(2022\)](#), the databases that include segmentation of the nucleus or cytoplasm of cervical cells are Herlev ([Jantzen et al. 2005](#)), ISBI2014 ([Ronneberger et al. 2015](#)) and ISBI2015 ([Lu et al. 2016](#)), BHS ([Araújo et al. 2019](#)), IRNet ([Zhou et al. 2019](#)) and Hybrid ([Zhu et al. 2021](#)).

Herlev is the oldest public cytology database, and its first publication dates from 2005. It contains segmentation of the nucleus and cytoplasm of 917 isolated cells obtained from Pap smear, classified into 7 classes, 3 normal and 4 abnormal, and its nomenclature does not follow TBS lesions nomenclature.

The ISBI database contains cervical cell images with groups of cells varying in number and degree of overlap. However, the images are synthetic and were artificially generated from liquid-based Extended Depth of Field (EDF) cervical cells segmentation.

Furthermore, ISBI has only normal cells and does not represent the pre-neoplastic cervical cell anomalies. ISBI 2014 and 2015 differ in the number of cells segmented for training and testing.

The Brazilian Health System (BHS) collection database has 194 cervical Pap smear images ([Araújo et al. 2019](#)). Among them, 108 images have at least one abnormal cell, and 86 images have only normal cells. It has five types of anomalies: squamous cell carcinoma (SC), high-grade squamous intraepithelial lesion (HSIL), low-grade squamous intraepithelial lesion (LSIL), atypical squamous cells of undetermined significance (ASC-US), and atypical squamous cells of undetermined significance, which cannot exclude high-grade lesions (ASC-H). The classification and cell bounding-box annotation were performed by three (3) specialist cytopathologists. In 2021, a complete classification database was delivered and named CRIC Cervix ([Rezende et al. 2021](#)).

The IRCNet and Hybrid databases will not be detailed as they are private, and their images and information are unavailable.

The limited number of cell segmentation databases interferes with the study and development of algorithms and computational tools that support medical decision-making systems in digital cytopathology since available data does not represent the routine and variability of cytology laboratories. Further data collection is required for the proper development of segmentation algorithms. Thus, this work will use and deliver a new segmentation database, CRIC Cervix-Seg, that includes nuclei and cytoplasm segmentation for real Pap smear images obtained from conventional cytology. CRIC Cervix-Seg used images from CRIC Cervix and was built using a web platform¹ for the labelling process. A complete description will be available in [Section 3.1.2](#).

[Table 1](#) presents a comparison between existing segmentation database and our proposal.

2.2. Nuclei segmentation using convolutional neural networks

Some studies in the literature have addressed the problem of cell segmentation using automatic methodologies with convolutional neural networks (CNNs). It is important to highlight that traditional machine learning methods ([Ushizima et al. 2014](#); [Lu et al. 2016](#)) are also used for the same purpose and whose results will be used for comparison ([Huang and Zhu 2020](#); [Win et al. 2020](#); [Wang et al. 2020](#)).

In 2015, [Ronneberger et al. \(2015\)](#) developed the U-net architecture, achieving the best results in the segmentation of

Table 1. Comparison of segmentation databases.

Property	CRIC Cervix-Seg	Herlev	SIPaKMeD	ISBI 2015	BHS
# Patches	-	-	-	-	194
#Contours/GT	3,200	917	4049	-	-
# Original images	400	917	966	945	194
Cells per image	Variable	1	Variable	Variable	Variable
Classified cells	3,200	917	4,049	-	-
Lesion type	6 class (TBS)	5 class	normal cells	normal cells	6 class (TBS)
Screening type	conventional	conventional	liquid-based	liquid-based	conventional
Validation	2 cytopathologists	2 cyto- 2 technicians	expert cyto- pathologists	cytopathologists	3 cytopathologists

biological images. The convolutional neural network consists of 28 (twenty-eight) layers of convolution, 9 convolutional blocks for *encoder* and *decoder*. The *encoder* is followed by *MaxPooling*, in which the most important information is selected for context capture. In the *decoder*, they are followed by transposed convolution layers to restore the image resolution and capture location information. Data augmentation techniques result in a good performance model based on the ISBI 2014 cell segmentation challenge.

Based on the contrast between the nucleus and cytoplasm, Gautam et al. (2018) proposed a 3-step methodology for segmenting nuclei from the Herlev database. A pre-processing selection of images with homogeneous chromatin patterns; the splits are trained separately on a model based on VGGNet. Finally, morphological postprocessing operations are applied to generate the final segmentation. The proposal surpassed previous results, with a recall of 0.91 and an f-score of 0.90.

Liu et al. (2018) presented a methodology for nuclei segmentation using the MaskRCNN model for prediction and *Fully connected conditional random field* (LFCCRF) for refining the edges of the results. The evaluation of the methods was carried out through the Herlev database and proved superior to previously presented methodologies.

Using the ResNeXt architecture (Xie et al. 2017) in tree topology and integrating pairs of adjacent layers, Zhang et al. (2019) proposed a new method to improve the capture of semantic and spatial information in the segmentation. The method obtained better results compared to the U-Net, FCN (Fully conventional network), and DeepLabv3+ (Chen et al. 2018) for precision, recall, and Dice coefficient in the ISBI2014 (Lu et al. 2015) database, and its private collection.

The main proposal of Araújo et al. (2019) was a deep learning tool to assist the cytopathologist in regions of interested ranking. This article was the first approach using the BHS database of conventional cytology, a subset of the CRIC Cervix database (Rezende et al. 2021). Center for Recognition and Inspection of Cells (CRIC) is a research consortium between researchers that aims to provide cell collections to the scientific community. Each image clump has a probability ranking of containing abnormal cells. A pre-trained LeNet architecture CNN applied a pixel-by-pixel classification of the probability of containing an abnormal cell. A post-segmentation process to remove noisy and non-cell structures was determined through the structuring element radius and a minimum area greater than the defined threshold – because of the applied parameter estimation process. With the final objective of targeting only abnormal cells, the results obtained 73% precision and MAP (*Mean Average Precision*) of 0.936 for classifying abnormal cells.

Using the Mask R-CNN architecture (He et al. 2018) pre-trained with the COCO database, Kurnianingsih et al. (2019) addressed cell segmentation and classification in a more compact VGG architecture with seven smaller layers and filters. Furthermore, geometric transformations to increase data in the Herlev database made it possible to reach an average precision of 0.92 for segmenting the cellular structure.

Precise segmentation of overlapping cells was the main factor considered by Wan et al. (2019) when proposing the methodology aimed at CNNs. The first step was to use the TernausNet (Iglovikov and Shvets 2018) network based on the

U-Net architecture with the weights of the VGG11 pre-trained by the ImageNet database, with the last fully connected layer giving rise to a convolution layer with 512 filters to sort between the nucleus, cytoplasm, and background. Post-processing was used with the Adaboost classifier and features extraction, Oriented Gradient Histogram (HOG), and the GLSZM descriptor (*Gray Level Size Zone Matrix*). Based on the position of the nucleus concerning the cytoplasm, an ROI (region of interest) was delimited. The cytoplasm was detected using DeepLabV2 (Chen et al. 2017) model and refined by CRF (*Conditional random fields*) and DRLSE (*Distance regularised level set evolution*). Furthermore, through the use of ISBI2014, ISBI2015, and INHouse databases, they suggest a 15% improvement in the Dice coefficient compared to previous works.

Using the FC-DenseNet56 model, Kůrková et al. (2020) addressed the importance of nuclei segmentation for the automation of cervical cell analysis. The model was trained on the Herlev database with an accuracy of 0.99, and the result was superior to previous works. Furthermore, the work reached 0.93 of the Dice coefficient, which characterises a segmentation close to the ground truth.

In 2020, Zou et al. (2020) presented a convolutional neural network model for segmenting and further classifying cells in a private database. The approach consists of segmenting the nuclei using the Mask R-CNN model. In addition, through 15 architectures of CNNs pre-trained with the ImageNet database, the classification centred on the nucleus of the newly segmented cell was obtained. Finally, the authors conclude that all models obtained good results, with the approach based on nuclei being a good methodology for classification proposals.

Based on the effectiveness of FCNs, Hussain et al. (2020) presented an approach to cervical nuclei segmentation. The methodology consists of the pre-trained SRM model (*Shape representation model*) to learn the information from the nuclei based on the ground truth. The model output is used as input to the proposed architecture. Based on the U-net model, the architecture consists of the *encoder-decoder* methodology, with only the addition of residual connections and densely connected blocks. The proposed method obtained the best results in the Dice, precision, recall, and accuracy measures compared to the U-net and Mask R-CNN models for Herlev and a private database.

Using a model similar to U-net, Yang et al. (2020) introduced a methodology that consists of using the ICM (*Interacting convolutional module*) and IPRCM (*Internal*

pyramid resolution complementing module) methods. ICM is a convolutional module that adds new details through 2 filter sizes. The IPRCM was applied to combine feature maps of different resolutions, considering that the information would complement each other. With an accuracy of 0.972, this work outperformed previous results for the Herlev database.

Some authors used databases containing a single (Herlev) or synthetic cells (ISBI2014 and ISBI2015). Therefore, they do not refer to real situations with a higher level of complexity in terms of cell overlap and artefacts in the image. There are also authors using liquid-based cytology in the databases (ISBI2014, ISBI2015, SIPAKMEed, and INHouse). Although they represent real images, the liquid-based exam is more expensive than the conventional one.

**Table 2.** Summary of segmentation related works.

Articles	Database	Segmentation method
Ronneberger et al. (2015)	ISBI2014	U-net neural network
Gautam et al. (2018)	Herlev	VGGNet based model
Liu et al. (2018)	Herlev	Mask-RCNN and LFCCRF model
Zhang et al. (2019)	ISBI 2014 and private collection	Architecture with network topology based at ResNeXt and TFA algorithm
Araújo et al. (2019)	BHS	LeNet neural network
Kurnianingsih et al. (2019)	Herlev database	Mask R-CNN
Wan et al. (2019)	ISBI2014, ISBI2015 and INHouse	TernausNet
Kůrková et al. (2020)	Herlev	FC-DenseNet56 model
Zou et al. (2020)	Private - liquid based cytology	Mask R-CNN
Hussain et al. (2020)	Herlev and private database	U-net based architecture
Huang et al. (2020)	ISBI2014 and ISBI15	FCM
Yang et al. (2020)	Herlev	ICM and IPRCM
Huang and Zhu (2020)	ISBI2014	GGCMM
Win et al. (2020)	SIPaKMeD and Herlev	<i>Watershed Transformation</i>
Wang et al. (2020)	ISBI2015	FCM and ADLS

The authors Win et al. (2020) and (Hussain et al. 2020) used real conventional bases: SIPaKMeD and a private database, respectively. The SIPaKMeD database, although presenting real images, is not classified using the Bethesda system of conventional cytology for diagnosing cell lesions. The database used by Hussain et al. (2020) comprises three databases – Herlev, liquid, and conventional. Although one of the databases used is conventional, the authors did not demonstrate the results for each database separately. Thus, evaluating the method's performance in conventional images was impossible.

Finally, (Huang and Zhu 2020) and (Wang et al. 2020) also used ISBI2014 and ISBI2015, but with conventional machine learning methodologies instead of neural network proposals.

Our proposal focuses on the automatic segmentation of cervical nuclei obtained from Pap smears. Unlike related works, we propose to use a real database with multiple cells to compare U-Net and Segnet convolutional neural networks. Such images are similar to those found in the routine of cytology laboratories and may be useful for computer aided decision (CAD) systems. Table 2 summarises the related works covered in this section. The first column presents the authors, followed by the databases and the proposed segmentation method.

3. Materials and methods

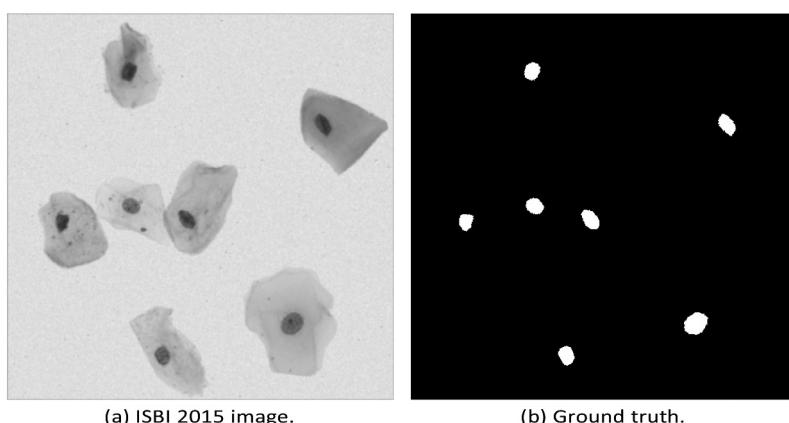
This section presents the used databases, the two CNNs architectures used to segment the nuclei in cervical cells, the U-net (Ronneberger et al. 2015), and SegNet (Kendall et al. 2015) networks, and the evaluation measures. We investigate the use of convolutional neural networks to segment cervical nuclei in images similar to those of a Pap smear, typical images observed by a cytopathologist.

3.1. Multiple cells databases

The objective is to investigate different cell segmentation proposals in images with a group of cells similar to those obtained in pap smears. The experiments will be done first in the synthetic database (ISBI2015 - (Lu et al. 2016)), then in the real database (CRIC Cervix-Seg), and the results will be compared considering the different methodologies.

3.1.1. ISBI

The ISBI2015 database comprises Pap smear images, 16 EDF real cervical cytology images, and 945 synthetic images. The training database has 45 synthetic images, 90 for validation and 810 for testing. Figure 1 presents an example of the original image and nuclei ground truth. The synthetic images were

**Figure 1.** Nuclei Segmentation in ISBI.

generated by segmenting the nuclei and cytoplasm of the EDF images through an overlay parameter ranging from 2 to 5 overlapping cells. As mentioned in [Section 2](#), the ISBI database has been widely used by researchers to compare results between different approaches.

3.1.2. CRIC Cervix-Seg

This paper proposed and delivered a new image database called *CRIC Cervix-Seg*, containing 3,200 segmented cellular nuclei and cytoplasm obtained from 400 real Pap smear images of CRIC Cervix (Rezende et al. 2021). The segmentations were obtained through the CRIC Searchable Image Database web platform², designed for manual annotation of cell structures, segmentation, and classification, being very useful in image labelling biological tools for machine learning algorithms. [Figure 2](#) presents examples of segmented nuclei and cytoplasm for different images. The images have only segmented some of the cells (nucleus and cytoplasm). Firstly, we would have an imbalance between cell types, as there are many more normal cells, and secondly, due to the difficulty of performing all the manual segmentation in an image. The database stands out from other proposals for including a diversity of cells and lesions in the segmentation approach: negative for intraepithelial lesion or malignancy (NILM); atypical squamous cells of undetermined significance, possibly non-neoplastic (ASC-US); low-grade squamous intraepithelial lesion (LSIL); atypical squamous cells which cannot exclude high-grade lesions (ASC-H); high-grade squamous intraepithelial lesion (HSIL); and squamous cell carcinoma (SC).

This database is the first to cover conventional cytology that follows internationally adopted protocols, Bethesda System, for evaluating cellular lesions. Another contribution is that it provides a broad diversity of neoplastic lesions with a variable number of cells, including overlap and inflammatory agents, never reported in the literature.

Papanicolaou staining is commonly performed for the exams used in this study. It is a routine staining in cytopathology laboratories and requires no additional expenditure. The cell nucleus is specifically stained with Harris haematoxylin, a dye widely used in pathology, which makes up the set of dyes for the previously mentioned staining.

The photo-documentation was performed through conventional bright field microscopy with a 40x objective and a 10x eyepiece, using a Zeiss AxionCam MRc digital camera coupled

to the Zeiss AxioImager Z2 microscope with the Axion Vision Zeiss Software, which is at the Microscopy facility of the Biological Sciences Research Center at Federal University of Ouro Preto.

The segmentation collection is performed by two members of the Cytology Laboratory team, with experience of eleven (11) and twenty (20) years working in diagnostic cytology. Each cell was manually segmented following defined criteria for delimitation of the two regions. The process of marking the contour (x,y) of the nucleus is followed by the contour of the cytoplasm. After marking both contours, a classification was assigned to the cell following morphological criteria prescribed by the nomenclature of the Bethesda System. A second pathologist reviewed the segmentation and classification markings; if their answers agreed, the information was approved. Otherwise, a consensus was reached between the two cytologists to define the regions and classification.

3.1.2.1. Pre-processing.

As mentioned, each image in the CRIC Cervix-Seg database does not have all segmented nuclei and cytoplasm. However, neural network training requires cell segmentation ground truth masks. Thus, we proposed to cut regions in the image for neural training and testing. We use sub-images of 128×128 that contain segmented nuclei. The cuttings of the adopted sub-images are based on the available GPU memory for neural network experiments.

The sub-images contain information regarding the position (x,y) of the contour of cell nuclei, [Figure 3\(a\)](#). Instead of using it to build a central bounding box, we generate sub-images (grid boxes) that avoid the centralised nuclei training bias, resulting in images with nuclei in different positions, green boxes in [Figure 3\(b\)](#).

So, the pre-processing methodology for training followed the steps: (i) division of database images into sub-images of 128×128 , (ii) visual and manual inspection of sub-images to the cytopathologist for approval, keeping just sub-images that have segmented nuclei, green bounding boxes in [Figure 3\(b\)](#), (iii) ground truth mask production, and, finally, (iv) storage of sub-images and their segmentation masks.

The segmentation database results in 3,200 sub-images with segmented nuclei, resulting in data with a diversity of cellular lesions. The cropped images were randomly divided into training (training and validation) and testing; more details are in [Section 3.3](#).

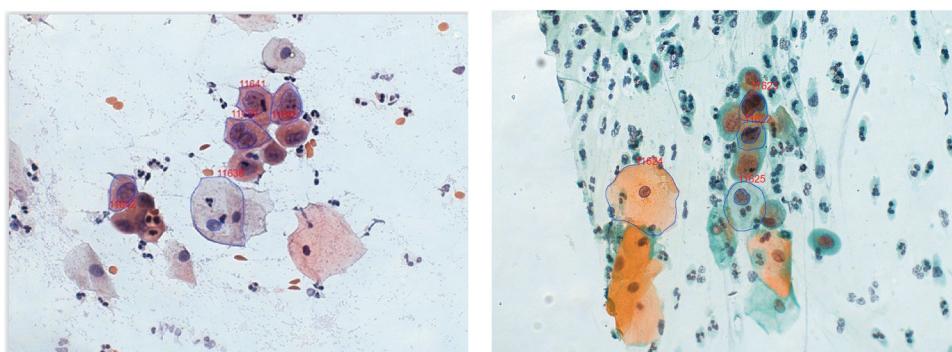


Figure 2. CRIC Cervix-Seg example images.

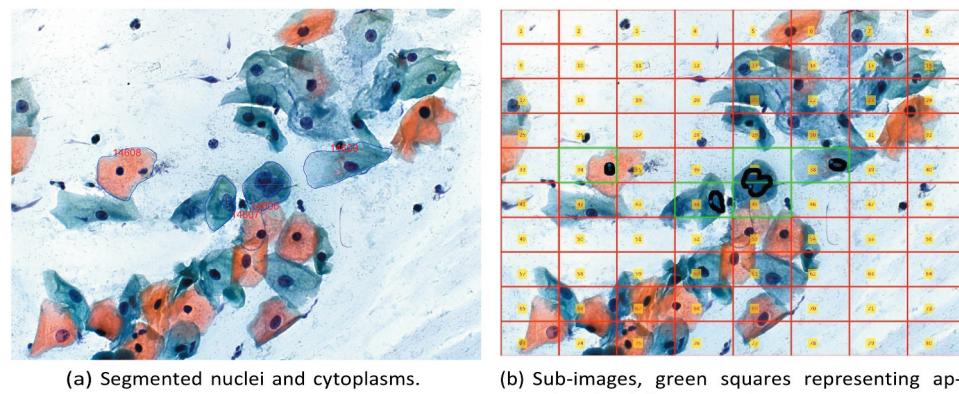


Figure 3. Segmentation database images, CRIC Cervix-Seg, original image and sub-images.

The segmentation database is part of a project of open availability of data and codes carried out by CRIC researchers. Part of the information is already available on the website and was published by Rezende et al. (2021), which is related to the detection and classification of 11,534 cervical cells. Segmentation data will be available on the platform once the article is published.

The database will advance developing methodologies and/or automated products in computer-assisted cytology regarding classification and segmentation. It will benefit machine learning researchers who need labelled data to test algorithm performance, particularly those related to deep convolutional neural networks. It will allow discoveries about computational and biological descriptors, subsidising the creation of new hypotheses about the differentiation of cell lesions and also help researchers in the biological area.

Concerning ethical issues, the Research Ethics Committee approved the work at the Federal University of Ouro Preto, Minas Gerais, Brazil, through protocol number 1,944,523. The Research Ethics Committee, in the protocol agreements, waived the requirement to obtain consent from participants for data sharing. Papanicolaou samples were obtained from female patients in the Southeast region of Brazil, South America. Pap smears were processed and analysed at Cytology Laboratory of the Faculty of Pharmacy at the Federal University of Ouro Preto, Minas Gerais, Brazil.

3.2. Convolutional neural network architectures

This section presents two convolutional neural networks to segment nuclei obtained in Pap smear images: the Unet and the Segnet. These CNNs are two neural networks used with great success in medical image segmentation once the encoder-decoder architectures allow good boundary delineation in

segmented structures. So far, the main cervical cell segmentation results used Unet or Unet-based networks (Araújo et al. 2019; Huang et al. 2020; Huang and Zhu 2020; Wang et al. 2020). Table 3 presents a summary of architectures used for segmentation.

In Segnet, only the pooling indices are transferred to the expansion path from the compression path, using less memory. In UNet, entire feature maps are transferred from compression path to expansion path making, using a lot of memory. Thus, SegNet is less time consuming, but its disadvantage is that it tends to lose neighbouring information when unpooling from low-resolution feature maps.

The evaluations will be carried out initially on cropped images for both networks and then on images of grouped cells, typically observed by a pathologist in a pap smear.

3.2.1. Unet

The U-net architecture (Ronneberger et al. 2015) presented for the segmentation of biomedical images in the ISBI2015 segmentation challenge was highlighted by the approach of reusing the feature maps of the *encoder* in the *decoder* of the network. This fully convolutional architecture was suitable for semantic segmentation. The features were extracted in the encoder, and the spatial dimension decreased proportions. Each convolution block consists of 3 convolution layers followed by the activation function, ending the block with a MaxPooling operation, see Figure 4. The decoder consists of convolution blocks followed by transposed convolutions to restore image resolution and capture location data.

The copy-and-crop approach proposed by U-Net consists of clipping the feature maps to the input dimension of the layers that reconstruct the original aspect ratio of the image. In this way, a U-net reuses the encoder outputs as input decoder, contributing to a new set of features.

Table 3. Summary of architecture characteristics.

Neural network	Number of layers	Filter layers
U-net	28 layers	9 encoder (convolution, batch normalization + dropout/maxpooling) 9 decoder (convolution, batch normalization+transposed convolution layers)
Segnet	36 layers	18 encoder (convolution, batch normalization, RELU and pooling) 18 decoder (convolution, upsampling, convolution)

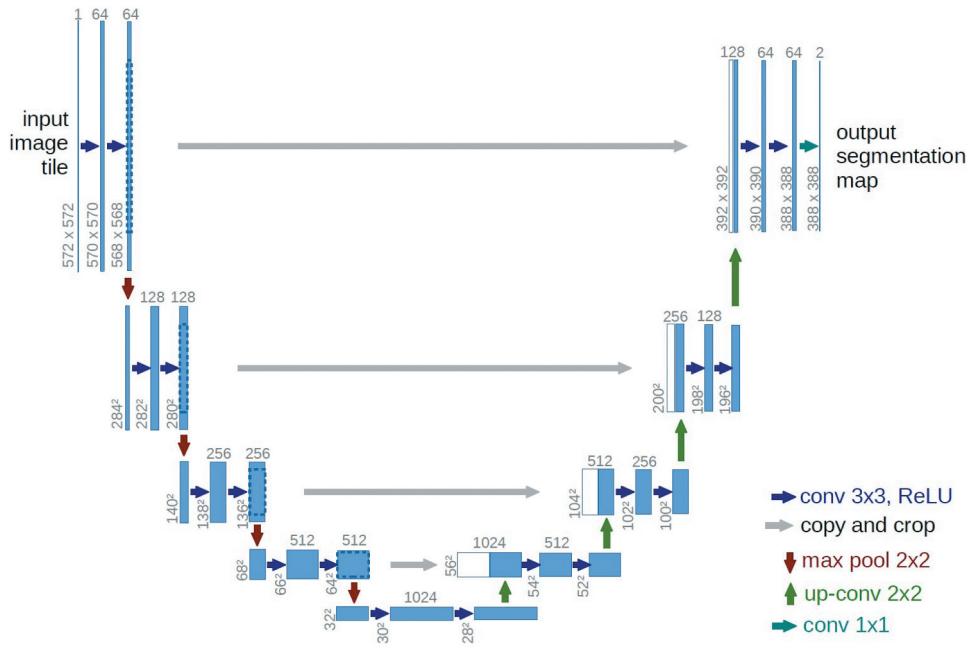


Figure 4. U-net architecture representation, proposed by Ronneberger et al. (2015).

3.2.2. Segnet

Although no works in the literature involve SegNet for the segmentation of cervical cytology nuclei, the architecture excelled in retaining details through the street image dataset by Kendall et al. (2015). Unlike U-net, the SegNet architecture brought the novelty of reconstructing feature maps through stored subsampling indices and their positions. In this way, the

positions of the most significant *pixels* of the feature maps are maintained so that the oversampling can reconstruct the original resolution of the image. The technique is demonstrated in Figure 5.

The encoder consists of a VGG16 architecture pre-trained by the ImageNet database, in which we use the transfer learning technique. Removing the fully connected layers of the original

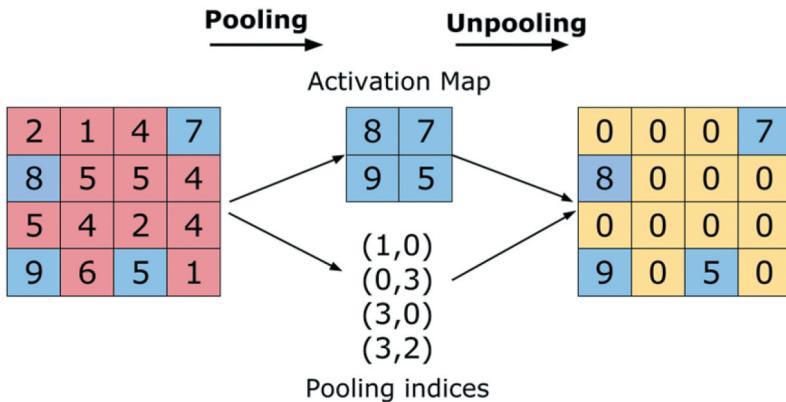


Figure 5. Reconstruction technique using *MaxPooling* indices.

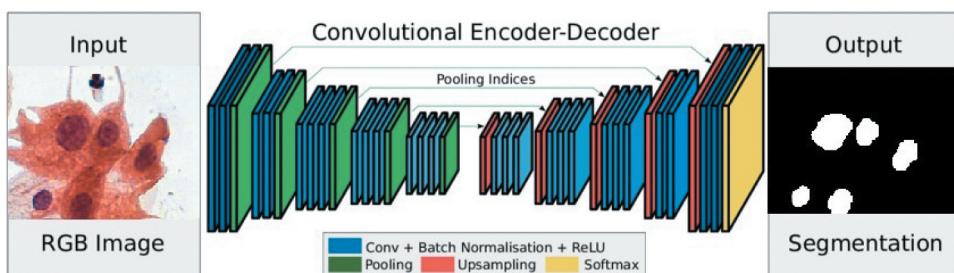


Figure 6. SegNet architecture adapted by Kendall et al. (2015).



architecture decreased the number of parameters from 134 M to 14.7 M. In addition, each convolution block was followed by batch normalisation, activation function, and MaxPooling, see Figure 6.

In Segnet, only the pooling indices are transferred to the expansion path from the compression path, using less memory. In UNet, entire feature maps are transferred from compression path to expansion path making, using a lot of memory. Thus, SegNet is less time consuming, but its disadvantage is that it tends to lose neighbouring information when unpooling from low-resolution feature maps, the 128×128 .

3.3. Experiments

The training of the convolutional neural networks impacts the model generalisation because of the difficulty in each sample in dividing the sets. For segmentation, it is common to split the data randomly.

The CRIC Cervix-Seg dataset (sub-images of 128×128) was randomly divided into 80% for training and 20% for testing. Finally, the result presented is an average of the results of each experiment. ISBI2015 is available with the division.

Another important aspect is that convolutional neural networks have many parameters to optimise. Thus, they require large data for model learning, as small sets may not achieve desirable generalisation power. In these cases, the model is overfitted when the model adapts well to the training but performs poorly for the test data. To minimise overfitting, we use data augmentation, a technique that generates new training data (Goodfellow et al. 2016).

The used data augmentation method includes geometric transformations to generate other images. To increase the two databases, it is necessary to consider the impact on morphological characteristics of the cell nucleus, in which changes can impact the diagnosis. In this work, rotation, inversion (horizontal and vertical), and modification operations were performed in the brightness and contrast of the image according to a set of modification parameters.

In addition to data augmentation, we incorporated other strategies to minimise overfitting and build a more accurate model representing the dataset, such as dropout, batch normalisation, and cross-validation.

In the SegNet architecture, the random initialisation of weights was avoided by reusing the weights of pre-trained architectures. In this sense, the Transfer of learning technique presented in Section 3.4.1 was applied due to its encoder being composed of a VGG16 (Simonyan and Zisserman 2015).

The training of both networks was divided into training cycles. After the model completes a training epoch, it is verified whether or not there was an improvement concerning the result of the previous epoch. If the validation loss decreases, the network weights are saved. Ensuring the values of the trainable parameters are the best found during training. Also, the application of the early stop technique determines that the

training cycle is interrupted after a number π_{pa} of epochs without improvement.

The model will continue with the training cycle until the number of epochs without improvement and the upper limit of epochs is not reached. As soon as the cycle is interrupted, a check is made whether the lower limit π_{lr} of the learning rate (*Learning rate*) has been reached. If true, the training is terminated; otherwise, the learning rate decreases. Such operation improves the network prediction accuracy since the learning rate is responsible for the step size in the gradient descent. Also, according to You et al. (2019), starting with a large learning rate prevents the network from memorising noisy data while gradually decreasing it improves the learning of complex patterns but takes more processing time.

3.3.1. Transfer learning

Compared to traditional machine learning methods, deep learning requires large-scale data for training. Building a database with thousands of samples is a highly complex task, in addition to being susceptible to the high cost of labelling (Tan et al. 2018).

Transfer Learning leverages previous network weights to facilitate predictive modelling (Lu et al. 2015). First, the network is trained with a labelled large-scale database. Then, the parameter values of the inner layers are transferred to a new network. To compensate for the difference between the two databases, adaptation layers are added to enable the extraction of features from the new dataset (Oquab et al. 2014).

Deep Fine Tuning (DFT) was used to refine the parameters of the CNNs. Although it has a higher computational cost, it can benefit applications where the problem domain differs from that used in network training. For both networks, we used the ImageNet dataset.

3.3.2. Evaluation metrics

According to Matias et al. (2021), among the most recurrent metrics for the segmentation of Papanicolaou smears in the period 2016–2020 are accuracy and Dice coefficient, followed by recall and precision.

The metric compares common pixels from two regions of interest (ROIs). The formula equals twice the number of elements common to both images (classified correctly) divided by the sum of the total number of pixels in each one. The similarity calculation compares regions segmented by our proposal and Ground Truth (GT). The performance of individual nuclei segmentation is assessed using the average Dice Coefficient, Equation (1).

$$\text{Dice} = \frac{2TP}{2TP + FP + FN} \quad (1)$$

The number of true positives (TP) indicates the pixels correctly classified as belonging to a region. True negatives (TN) denote the pixels not belonging to the nuclei correctly classified as background. False positives (FP) refer to pixels in the background wrongly classified as part of the region, and false

negatives (FN), pixels belonging to the nuclei wrongly classified as belonging to the background.

Accuracy indicates the overall performance of the model, the percentage of true negative (TN) and true positive (TP) out of the entire rating in the dataset, Equation (2).

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

The recall is the percentage of positive data correctly classified as positive. Also, it is convenient in hypotheses where false negatives are more harmful than false positives. It is defined by Equation (3).

$$Revoc = \frac{TP}{TP + FN} \quad (3)$$

Precision exposes the positive ratings of the result of our proposal and how many are correct. It is a favourable option when false positives are more harmful than false negatives. It is defined by Equation (4).

$$Prec = \frac{TP}{TP + FP} \quad (4)$$

4. Results and discussion

We implemented the U-net and SegNet networks, and the experiments were performed for the ISBI2015 and CRIC Cervix-Seg databases. The experiments run through the *Google Collaboratory*³ virtual machine. The construction of the CNNs was done in *Python* using the library *Tensorflow/Keras*⁴.

The network parameters were: *batch size*, *activation function*, and *batch normalisation*. The batch size varied according to the limitations of the *hardware*, varied from 16, 32, and 64. The application of batch normalisation evaluates the generalisation capacity of the technique and whether it can be used or not. Finally, the activation functions prevent neuron death caused by the ReLU function application used in the original architectures. Table 4 presents the variation of parameters for training.

Regarding the execution time of neural networks, each parameter configuration results in different execution times.

Table 4. Network parameters during training.

Parameters	Values
Batch size	[16, 32, 64]
Batch normalization	[True, False]
Activation function	[ReLU, Mish, Swish, Elu, LeakyRelu]

Considering the best segmentation results, *U-net without batch normalisation*, the average time for training was around 30 hours.

All experiments were performed four times over 300 epochs (or training cycles). The initial learning rate was obtained by observing the training of architectures in an interval of $[10^{-7}, 10^{-2}]$ in 300 epochs. Thus, the learning rate was set to 10^{-3} and divided by 10 when necessary until reaching the lower bound of $\pi_{lr} = 10^{-6}$. We proposed an *early stop* that occurred in the training cycle when the model reached $\pi_{pa} = 25$ epochs without improvement.

The predictions were obtained through the activation function *sigmoid* in a probability interval from 0 to 1. For the execution of the experiments, k distinct subsets of the databases were generated, with $k\text{-fold} = 10$. Then, these partitions were trained and evaluated separately. Finally, the result presented is an average of the results of each experiment, for each experiment. Table 5 presents the results of U-net and SegNet for segmenting nuclei from the ISBI2015 database.

The best Dice results were obtained for the RELU activation function in different batch sizes. SegNet method without batch normalisation obtained the best results for precision, 0.9933, and Dice coefficient, 0.9783. In the U-net prediction without batch normalisation, Figure 7, the model presented more false negatives by classifying nuclei as background. Both methodologies, U-net and SegNet, surpassed previous works for the Dice coefficient, as shown in Table 5.

Table 6 presents the results for CRIC Cervix-Seg database for the considered architectures. The RELU in different batch sizes was the best activation function. The *U-net without batch normalisation* obtained the best segmentation results, 0.9524 for precision, 0.9336 for recall, and 0.9429 for the Dice coefficient. The best accuracy value was also for a U-net (with batch normalisation). Overall, U-net was slightly superior to SegNet, able to segment cervical cell nuclei with quality. As CRIC Cervix-Seg is an unpublished database, there are no other works in the literature.

In the prediction results, the colour variations indicate the possibility of the *pixel* belonging to the nuclei region, ranging from yellow to purple (through blue), where yellow indicates more likely and purple less likely. Figures 8 present the original sub-images in (a), the GTs in (b), and its predictions in (c) for SegNet without batch normalisation. Figures 9, 10 and 11 illustrate results for SegNet with batch normalisation, U-net without and with batch normalisation, respectively.

Table 5. Results for nuclei segmentation using ISBI2015 database.

Articles	Segmentation method	Precision	Recall	Dice	Accuracy
Ushizima et al. (2014)	Super-pixels and Voronoi diagram	0.9590	0.8950	0.9140	-
Saha et al. (2017)	CiscFC	0.9680	0.8820	0.9380	-
Phouladhy et al. (2017)	Iterative method of binarization	0.9610	0.9330	0.9380	-
Huang et al. (2020)	Multi-scale fuzzy clustering	0.9810	0.9390	0.9360	-
Hoque et al. (2021)	Contour based method	0.9780	0.9333	-	-
Our proposal	U-net with batch normalization	0.9794	0.9714	0.9743	0.9957
Our proposal	U-net without batch normalization	0.9767	0.9603	0.9684	0.9948
Our proposal	SegNet with batch normalization	0.9725	0.9706	0.9726	0.9961
Our proposal	SegNet without batch normalization	0.9933	0.9654	0.9783	0.9929

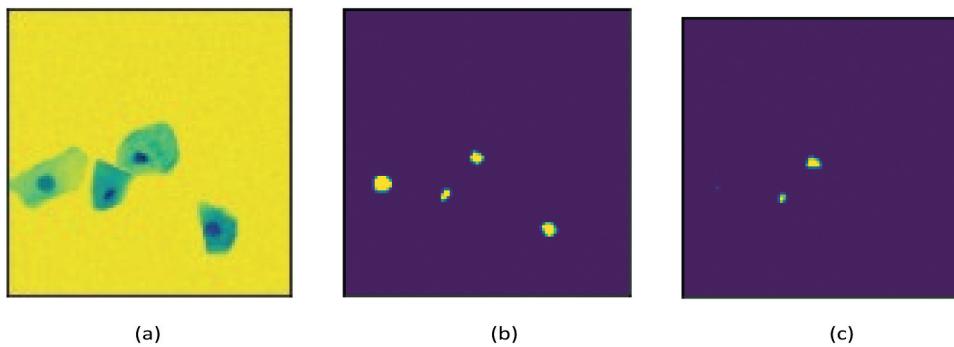


Figure 7. Prediction of U-net without batch normalization for ISBI2015.

Table 6. Results of the proposed methodology for nuclei segmentation using CRIC Cervix-Seg database.

Method	←Precision Recall Dice→			Accuracy
	Precision	Recall	Dice	
U-net with batch normalization	0.9418	0.9307	0.9352	0.9582
U-net without batch normalization	0.9524	0.9336	0.9429	0.9560
SegNet with batch normalization	0.9027	0.9166	0.9096	0.9552
SegNet without batch normalization	0.8026	0.6544	0.7130	0.9042

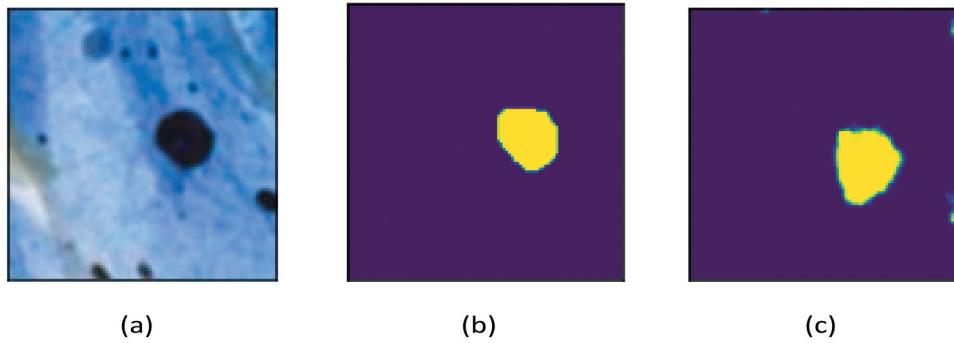


Figure 8. SegNet without batch normalization, original image, ground truth, and prediction.

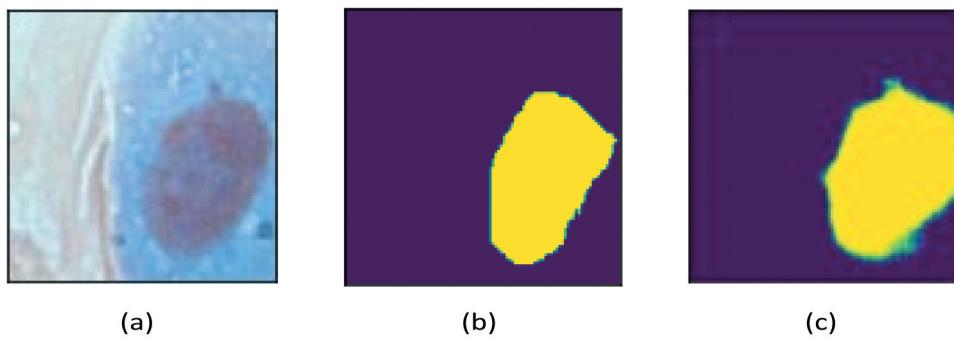


Figure 9. SegNet with batch normalization, original image, ground truth, and prediction.

Finally, we use the best parameter results from the U-net neural network to predict the results of nuclei segmentation in an image with multiple cells, nuclei, and other structures. The image is divided into grids of fixed size, analogous to the training pre-processing, each grid is presented to the network, and its results are processed individually and aggregated again

in the image. [Figure 12](#) (a) shows the original image, in (b) the GT, and (c) the prediction result as a function of pixel probabilities. In the prediction image, it is possible to observe that the nuclei close to or behind the neutrophils were partially or not segmented. The performance of the model generated results with a high number of true negatives. The model can segment

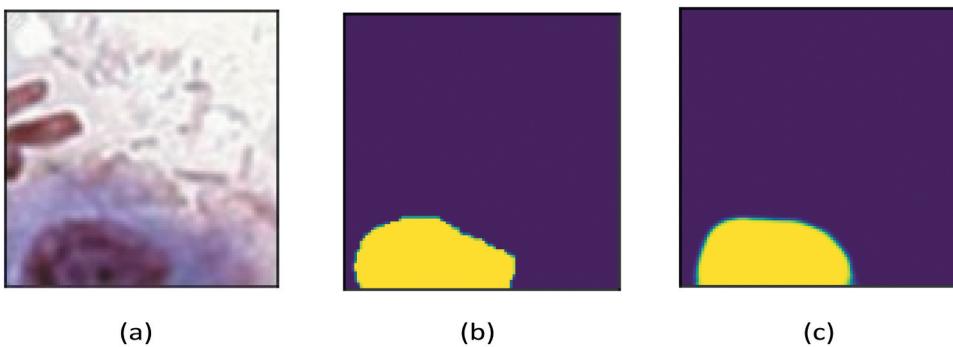


Figure 10. U-net without batch normalization, original image, ground truth and prediction.

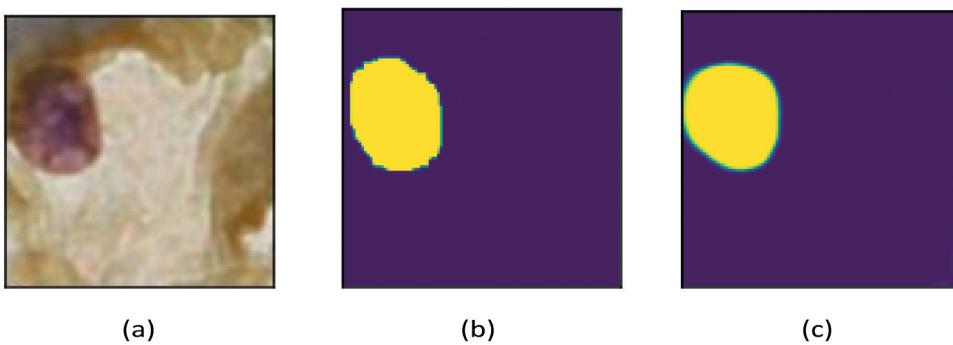


Figure 11. U-net with batch normalization, original image, ground truth and prediction.

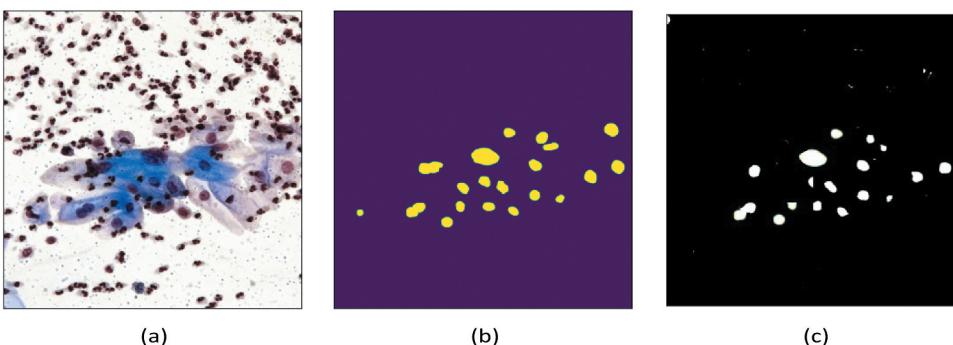


Figure 12. Image prediction for U-net neural network segmentation for CRIC Cervix-Seg.

most nuclei with quality, satisfactorily differentiating nuclei, background, and other structures.

5. Conclusions

This paper provided a methodology for automatically segmenting images of real cervical cells using convolutional neural networks, fine-tuning training, and a new database. The segmentation of cervical nuclei is an important step for classification since the nucleus presents essential information for the diagnosis and can be considered a pre-processing for automatic classification. Thus, this paper verified the performance of CNNs architectures when applied to the segmentation of real nuclei of conventional cervical cytology. The U-net and SegNet models were investigated through accuracy, Dice coefficient, recall, and precision measures.

It also highlights the publication of a segmentation database containing the nucleus and cytoplasm of cervical cells. To the best of our knowledge, our work is the first to use conventional cytology and images with different types of lesions classified by the Bethesda system. However, given the limited number of segmented cells, caution must be taken, such as dropout and batch normalisation, to avoid overfitting.

Finally, our findings indicated high precision, recall, and Dice coefficient; thus, the proposed methodology can segment real cervical nuclei, even in situations of overlap and different biological artefacts. Once new approaches and a segmentation model for real images are defined, future works will evolve into a computational segmentation tool followed by the classification and ranking of cervical nuclei, as well as submitting these results



to the evaluation of health experts and measuring the impact of this technology in the routine and support in the laboratories.

A very interesting result regarding segmentation was that in the ISBI 2015 database, which is well-behaved and synthetic, Segnet provided better Dice coefficient values.

Meanwhile, U-net obtained better results in real images with greater overlap and inflammatory agents. Such behaviour is consistent with Segnet characteristic of losing contour information during the decoder process. In future works, we intend to investigate this behaviour deeper since the images are more complex in cytology laboratory routines.

Notes

1. <https://database.cric.com.br/>.
2. <https://database.cric.com.br/>.
3. colab.research.google.com
4. www.tensorflow.org.

Acknowledgements

The authors would like to thank the Universidade Federal de Ouro Preto (UFOP), the Center for Recognition and Inspection of Cells (CRIC), the Group of Optimization and Algorithms (GOAL), the Microscopy facility of the Biological Sciences Research Center (NUPEB), and the eXtended Reality 4 Good (XR4GoodLab) for supporting this research.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) under Grant [Finance Code 001]; Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG under Grants [APQ-00751-19, APQ-01306-22, APQ-01518-21, PPM/CEX/FAPEMIG/676-17, and PPSUS-FAPEMIG/APQ-03740-17]; Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq under Grants [303266/2019-8, 305895/2019-2, 308947/2020-7]; Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação - PROPI/UFOP under Grants [19/2020, 23109.000928/2020-33, and 23109.000929/2020-88]; Ministry of Health under grant [905103/2020]; and Moore-Sloan Foundation, and Office of Science, of the U.S. Department of Energy [under the Contract No. DE-AC02-05CH11231]. The authors also acknowledge the Federal University of Ouro Preto (UFOP) and the Coordination for the Improvement of Higher Education Personnel (CAPES) through EDITAL PROPI/UFOP Nº 18/2022 for the support of the language proofreading service costs.

References

- Araújo FH, Silva RR, Ushizima DM, Rezende MT, Carneiro CM, Bianchi AGC, Medeiros FN. 2019. Deep learning for cell image segmentation and ranking. *Comput Med Imaging Graph*. 72:13–21. doi:[10.1016/j.compmedim.2019.01.003](https://doi.org/10.1016/j.compmedim.2019.01.003).
- Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. 2017. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans Pattern Anal Mach Intell*. 40(4):834–848. doi:[10.1109/TPAMI.2017.2699184](https://doi.org/10.1109/TPAMI.2017.2699184).
- Chisale Mabotja M, Levin J, Kawonga M, Grce M. 2021. Beliefs and perceptions regarding cervical cancer and screening associated with pap smear uptake in johannesburg: a cross-sectional study. *Plos One*. 16(2):e0246574. doi:[10.1371/journal.pone.0246574](https://doi.org/10.1371/journal.pone.0246574).
- Diniz DN, Vitor RF, Bianchi AG, Delabrida S, Carneiro CM, Ushizima DM, de Medeiros FN, Souza MJ. 2021. An ensemble method for nuclei detection of overlapping cervical cells. *Expert Syst Appl*. 185:115642. doi:[10.1016/j.eswa.2021.115642](https://doi.org/10.1016/j.eswa.2021.115642).
- Gautam S, Bhavsar A, Sao AK, H KK. 2018. CNN based segmentation of nuclei in PAP-smear images with selective pre-processing. In: Tomaszewski J Gurcan MeditorsMedical imaging 2018: digital Pathology Vol. 10581 February 10-15 Houston, USAInternational Society for Optics and Photonics; SPIE p. 105810X
- Goodfellow I, Bengio Y, Courville A. 2016. Deep learning. Cambridge, MA: MIT press .
- He K, Gkioxari G, Dollár P, Girshick RB. 2018. Mask R-CNN; [arXiv:1703.06870v3 [cs.CV]]; January. Available from: <https://arxiv.org/pdf/1703.06870.pdf>.
- Hoque IT, Ibtehaz N, Chakravarty S, Rahman MS, Rahman MS. 2021. A contour property based approach to segment nuclei in cervical cytology images. *BMC Med Imaging*. 21(1):1–12. doi:[10.1186/s12880-020-00533-9](https://doi.org/10.1186/s12880-020-00533-9).
- Huang J, Wang T, Zheng D, He Y. 2020. Nucleus segmentation of cervical cytology images based on multi-scale fuzzy clustering algorithm. *Bioengineered*. 11(1):484–501. doi:[10.1080/21655979.2020.1747834](https://doi.org/10.1080/21655979.2020.1747834).
- Huang Y, Zhu H. 2020. Segmentation of overlapped cervical cells using asymmetric mixture model and shape constraint level set method. *Math Probl Eng*. 2020:3728572. doi:[10.1155/2020/3728572](https://doi.org/10.1155/2020/3728572).
- Hussain E, Mahanta LB, Das CR, Choudhury M, Chowdhury M. 2020. A shape context fully convolutional neural network for segmentation and classification of cervical nuclei in pap smear images. *Artificial Intelligence in Medicine*:101897.
- Iglorikov V, Shvets A. 2018. Ternausnet: u-net with VGG11 encoder pre-trained on imagenet for image segmentation; [arXiv:1801.05746v1 [cs.CV]]; January. Available from: <https://arxiv.org/pdf/1801.05746.pdf>.
- Jantzen J, Norup J, Dounias G, Bjerregaard B. 2005. Pap-smear benchmark data for pattern classification. In: Proceedings of the Nature-Inspired Smart Information Systems Symposium (NiSIS 2005); October 4-5; Albufeira, Portugal. NiSIS. p. 1–9. Available from: <https://bitly.com/PGIElbc>.
- Jiang H, Zhou Y, Lin Y, Chan RC, Liu J, Chen H. 2022. Deep learning for computational cytology: a survey; [arXiv:2202.05126v2 [eess.IV]]; February. Available from: <https://arxiv.org/pdf/2202.05126.pdf>.
- Kendall A, Badrinarayanan V, Cipolla R. 2015. Bayesian segnet: model uncertainty in deep convolutional encoder-decoder architectures for scene understanding; [arXiv:1511.02680v2 [cs.CV]]; October. arXiv:1511.02680v2 [cs.CV]: <https://arxiv.org/pdf/1511.02680.pdf>.
- Kůrková V, Manolopoulos Y, Hammer B, Iliadis, L, Maglogiannis, I. 2020. A Survey on Deep Transfer Learning. *Artificial Neural Networks and Machine Learning – ICANN 2018. Lecture Notes in Computer Science*. Cham: IEEE. p. 1274–1278. https://link.springer.com/chapter/10.1007/978-3-030-01424-7_27
- Kurnianingsih, Allehaibi KHS, Nugroho LE, Widyawan, Lazuardi L, Prabuwono AS, Mantoro T. 2019. Segmentation and classification of cervical cells using deep learning. *IEEE Access*. 7:116925–116941. doi:[10.1109/ACCESS.2019.2936017](https://doi.org/10.1109/ACCESS.2019.2936017).
- LC Chen, Zhu Y, Papandreou G, Schroff F, Adam H. 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: Ferrari V, Hebert M, Sminchisescu C, and Weiss Y, editors, *Computer Vision – ECCV 2018*. (Vol. 11211, pp. 833–851). Lecture Notes in Computer Science, Cham: Springer. https://link.springer.com/chapter/10.1007/978-3-030-01234-2_49
- Liang LA, Einzmann T, Franzen A, Schwarzer K, Schauberger G, Schriefer D, Radde K, Zeissig SR, Ikenberg H, Meijer CJ, et al. 2021. Cervical cancer screening: comparison of conventional pap smear test, liquid-based cytology, and human papillomavirus testing as stand-alone or cotesting strategies. *Cancer Epidemiol Prev Biomarkers*. 30(3):474–484. doi:[10.1158/1055-9965.EPI-20-1003](https://doi.org/10.1158/1055-9965.EPI-20-1003).
- Liu Y, Zhang P, Song Q, Li A, Zhang P, Gui Z. 2018. Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. *IEEE Access*. 6:53709–53721. doi:[10.1109/ACCESS.2018.2871153](https://doi.org/10.1109/ACCESS.2018.2871153).
- Lu Z, Carneiro G, Bradley AP. 2015. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical

- cells. *IEEE Trans Image Process*. 24(4):1261–1272. doi:[10.1109/TIP.2015.2389619](https://doi.org/10.1109/TIP.2015.2389619).
- Lu Z, Carneiro G, Bradley AP, Ushizima D, Nosrati MS, Bianchi AG, Carneiro CM, Hamarneh G. 2016. Evaluation of three algorithms for the segmentation of overlapping cervical cells. *IEEE J Biomed Health Inf*. 21(2):441–450. doi:[10.1109/JBHI.2016.2519686](https://doi.org/10.1109/JBHI.2016.2519686).
- Matias AV, Amorim JGA, Macarini LAB, Cerentini A, Onofre ASC, Onofre FBM, Daltoé FP, Stemmer MR, von Wangenheim A. 2021. What is the state of the art of computer vision-assisted cytology? A systematic literature review. *Comput Med Imaging Graph*. 91:101934. <https://www.sciencedirect.com/science/article/pii/S0895611121000835>
- Nayar R, Wilbur DC. 2017. The bethesda system for reporting cervical cytology: a historical perspective. *Acta Cytol*. 61(4–5):359–372. doi:[10.1159/000477556](https://doi.org/10.1159/000477556).
- Quab M, Bottou L, Laptev I, Sivic J. 2014. Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 23–28 June 2014, Columbus, OH, USA. p. 1717–1724. doi:[10.1109/CVPR.2014.222](https://doi.org/10.1109/CVPR.2014.222).
- Pangarkar MA. 2022. The bethesda system for reporting cervical cytology. *Cytjournal*. 19:28. doi:[10.25259/CMAS_03_07_2021](https://doi.org/10.25259/CMAS_03_07_2021).
- Phouladhy HA, Goldgof D, Hall LO, Mouton PR. 2017. A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images. *Comput Med Imaging Graph*. 59:38–49. doi:[10.1016/j.compmedimag.2017.06.007](https://doi.org/10.1016/j.compmedimag.2017.06.007).
- Rezende MT, Silva R, FdO B, Tobias AH, Oliveira PH, Machado TM, Costa CS, Medeiros FN, Ushizima DM, Carneiro CM, et al. 2021. Cric searchable image database as a public platform for conventional pap smear cytology data. *Sci Data*. 8(1):1–8. doi:[10.1038/s41597-021-00933-8](https://doi.org/10.1038/s41597-021-00933-8).
- Ronneberger O, Fischer P, Brox T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells W, and Frangi A, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Vol. 9351. Lecture Notes in Computer Science, Cham: Springer. pp. 234–241. https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28
- Saha R, Bajer M, Lee G. 2017. Circular shape constrained fuzzy clustering (ciscfc) for nucleus segmentation in pap smear images. *Comput Biol Med*. 85:13–23. doi:[10.1016/j.combiomed.2017.04.008](https://doi.org/10.1016/j.combiomed.2017.04.008).
- Simonyan K, Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition; [arXiv:1409.1556v6 [cs.CV]]; April. Available from: <https://arxiv.org/pdf/1409.1556.pdf>.
- Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. 2018. A survey on deep transfer learning. In: Kůrková V, Manolopoulos Y, Hammer B, Iliadis L, Maglogiannis I, editors. *International conference on artificial neural networks*. Cham: Springer; pp. 270–279. https://link.springer.com/chapter/10.1007/978-3-030-01424-7_27
- Ushizima DM, Bianchi AG, Carneiro CM. 2014. Segmentation of subcellular compartments combining superpixel representation with Voronoi diagrams. Lawrence Berkeley National Laboratory, Berkeley. Lawrence Berkeley National Laboratory.
- Wang T, Huang J, Zheng D, He Y. 2020. Nucleus segmentation of cervical cytology images based on depth information. *IEEE Access*. 8:75846–75859. doi:[10.1109/ACCESS.2020.2989369](https://doi.org/10.1109/ACCESS.2020.2989369).
- Wan T, Xu S, Sang C, Jin Y, Qin Z. 2019. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. *Neurocomputing*. 365:157–170. doi:[10.1016/j.neucom.2019.06.086](https://doi.org/10.1016/j.neucom.2019.06.086).
- WHO. 2022. World health organization - cervical cancer. Accessed = 2022-10-9; Available from: <https://www.who.int/news-room/fact-sheets/detail/cervical-cancer>.
- Win KP, Kitajadure Y, Hamamoto K, Myo Aung T. 2020. Computer-assisted screening for cervical cancer using digital image processing of pap smear images. *Appl Sci*. 10(5):1800. doi:[10.3390/app10051800](https://doi.org/10.3390/app10051800).
- Xie S, Girshick RB, Dollár P, Tu Z, He K. 2017. Aggregated residual transformations for deep neural networks; [arXiv:1611.05431v2 [cs.CV]]; April. Available from: <https://arxiv.org/pdf/1611.05431.pdf>.
- Yang X, Wu J, Yin Y. 2020. Interacting convolution with pyramid structure network for automated segmentation of cervical nuclei in pap smear images. In: Proceedings of the IEEE 17th International Symposium on Biomedical Imaging (ISBI 2020), 03–07 April 2020 Iowa City, IA, USA; April 3–7. IEEE. p. 499–502 <https://ieeexplore.ieee.org/document/9098383> doi:[10.1109/ISBI45749.2020.9098383](https://doi.org/10.1109/ISBI45749.2020.9098383).
- You K, Long M, Wang J, Jordan MI. 2019. How does learning rate decay help modern neural networks?; [arXiv:1908.01878v2 [cs.LG]]. September. Available from: <https://arxiv.org/pdf/1908.01878.pdf>.
- Zhang J, Liu Z, Du B, He J, Li G, Chen D. 2019. Binary tree-like network with two-path fusion attention feature for cervical cell nucleus segmentation. *Comput Biol Med*. 108:223–233. doi:[10.1016/j.combiomed.2019.03.011](https://doi.org/10.1016/j.combiomed.2019.03.011).
- Zhou Y, Chen H, Xu J, Dou Q, Heng PA. 2019. IRNet: Instance Relation Network for Overlapping Cervical Cell Segmentation . . In: Shen D, Liu T, Peters T, Staib L, Essert C, Zhou S, Yap P, and Khan A *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, editors ; vol. 11764 *Lecture Notes in Computer Science*. : Cham: Springer; p. 640–648 https://link.springer.com/chapter/10.1007/978-3-030-32239-7_71 doi:[10.1007/978-3-030-32239-7_71](https://doi.org/10.1007/978-3-030-32239-7_71).
- Zhu X, Li X, Ong K, Zhang W, Li W, Li L, Young D, Su Y, Shang B, Peng L, et al. 2021. Hybrid ai-assistive diagnostic model permits rapid tbs classification of cervical liquid-based thin-layer cell smears. *Nat Commun*. 12(1):1–12. doi:[10.1038/s41467-021-23913-3](https://doi.org/10.1038/s41467-021-23913-3).
- Zou J, Xue Z, Brown G, Long R, Antani S. 2020. Deep learning for nuclei segmentation and cell classification in cervical liquid based cytology. In: Chen P, and Deserno T *Editors Medical Imaging 2020: imaging Informatics for Healthcare, Research, and Applications Vol. 11318 SPIE Medical Imaging Houston, USASPIE Digital Library* p. 1131811 doi:[10.1117/12.2549547](https://doi.org/10.1117/12.2549547)

Data Preprocessing:

Load the microscopy images along with their corresponding labels from the provided dataset. Resize the images to a common size suitable for training the model, e.g., 224x224 pixels. (Optional: did not do) Normalize the pixel values of the images to a range between 0 and 1. Since there are no mask data provided, you will not be able to perform traditional segmentation-based preprocessing. However, you can still apply data augmentation techniques like rotation, flipping, and zooming to increase the diversity of your training data.

```
import cv2
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from tensorflow.keras.preprocessing.image import ImageDataGenerator
```

```
# Load the CSV file containing file paths and labels
csv_file = '../datasets/tb-wellgen-smear/v1/tb-labels.csv'
labels_df = pd.read_csv(csv_file)
```

```
# first 1000
labels_df = labels_df[:100]
labels_df
```

2024-03-19 23:55:38.393806: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use AVX2 AVX512F FMA, in other operations, rebuild TensorFlow with the appropriate compile flags
2024-03-19 23:55:39.312020: W tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT

	image	tb_positive	file_path
0	tb00000001.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
1	tb00000002.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
2	tb00000003.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
3	tb00000004.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
4	tb00000005.jpg	1	/home/ngsci/datasets/tb-wellgen-smear/images/0...
...
95	tb00000096.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
96	tb00000097.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
97	tb00000098.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
98	tb00000099.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...
99	tb00000100.jpg	0	/home/ngsci/datasets/tb-wellgen-smear/images/0...

100 rows × 3 columns

```
# Perform train-test split
train_df, test_df = train_test_split(labels_df, test_size=0.25, random_state=42)
```

```

# Define image dimensions
# height = 2048
# width = 2448
height = 32
width = 32

# Function to load and preprocess images
def preprocess_image(file_path):
    img = cv2.imread(file_path)
    img = cv2.resize(img, (height, width))
    img = img.astype(np.float32) / 255.0 # Normalize pixel values to [0,1] or mean and stdev instead
    return img

# Preprocess images for training with augmentation
train_datagen = ImageDataGenerator(
    rotation_range=20, # Random rotation up to 20 degrees
    width_shift_range=0.2, # Random shift up to 20% of the width
    height_shift_range=0.2, # Random shift up to 20% of the height
    shear_range=0.2, # Shear intensity
    zoom_range=0.2, # Random zoom up to 20%
    horizontal_flip=True, # Random horizontal flipping
    vertical_flip=True, # Random vertical flipping
    fill_mode='nearest' # Fill mode for points outside the input boundaries
)

train_images = np.array([preprocess_image(file_path) for file_path in train_df['file_path']])
train_labels = train_df['tb_positive'].values

# Initialize generator for training data
train_generator = train_datagen.flow(
    train_images,
    train_labels,
    batch_size=32,
    shuffle=True
)

# Preprocess images for testing without augmentation
test_images = np.array([preprocess_image(file_path) for file_path in test_df['file_path']])
test_labels = test_df['tb_positive'].values

num_classes=2
# Optionally, you can also retrieve class weights if dealing with imbalanced classes
class_weights = dict(zip(range(num_classes), (train_df['tb_positive'].value_counts() / len(train_df)).values))

```

Model Architecture:

The U-Net architecture consists of a contracting path to capture context followed by an expanding path for precise localization. It includes skip connections to concatenate features from the contracting path to the expanding path, aiding in better localization. You can design the U-Net model using convolutional layers, pooling layers for down-sampling, and upsampling layers for up-sampling. Since the task is binary classification (presence or absence of TB bacilli), you can use a binary cross-entropy loss function and a sigmoid activation function in the output layer. You can start with a pre-trained backbone like VGG or ResNet for feature extraction and then add the U-Net layers on top for segmentation.

```

from tensorflow.keras.models import Model
from tensorflow.keras.layers import Input, Conv2D, MaxPooling2D, Dropout, concatenate, Conv2DTranspose, Flatten, Dense

def unet_model(input_shape=(height, width, 3)):
    inputs = Input(input_shape)

    # Contracting path
    conv1 = Conv2D(64, 3, activation='relu', padding='same')(inputs)
    conv1 = Conv2D(64, 3, activation='relu', padding='same')(conv1)
    pool1 = MaxPooling2D(pool_size=(2, 2))(conv1)

    conv2 = Conv2D(128, 3, activation='relu', padding='same')(pool1)
    conv2 = Conv2D(128, 3, activation='relu', padding='same')(conv2)
    pool2 = MaxPooling2D(pool_size=(2, 2))(conv2)

    conv3 = Conv2D(256, 3, activation='relu', padding='same')(pool2)
    conv3 = Conv2D(256, 3, activation='relu', padding='same')(conv3)
    pool3 = MaxPooling2D(pool_size=(2, 2))(conv3)

    conv4 = Conv2D(512, 3, activation='relu', padding='same')(pool3)
    conv4 = Conv2D(512, 3, activation='relu', padding='same')(conv4)
    drop4 = Dropout(0.5)(conv4)
    pool4 = MaxPooling2D(pool_size=(2, 2))(drop4)

    # Bottom of U-Net
    conv5 = Conv2D(1024, 3, activation='relu', padding='same')(pool4)
    conv5 = Conv2D(1024, 3, activation='relu', padding='same')(conv5)
    drop5 = Dropout(0.5)(conv5)

    # Classification layer
    print("Shape of drop5:", drop5.shape) # Print the shape of drop5
    flatten = Flatten(input_shape=drop5.shape[1:])(drop5)
    dense1 = Dense(512, activation='relu')(flatten)
    dense2 = Dense(256, activation='relu')(dense1)
    outputs = Dense(1, activation='sigmoid')(dense2)

    model = Model(inputs=inputs, outputs=outputs)
    return model

```

```

# Compile the model
model = unet_model()
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

# Print model summary
model.summary()

```

🔗 Shape of drop5: (None, 2, 2, 1024)
Model: "model"

Layer (type)	Output Shape	Param #
<hr/>		
input_1 (InputLayer)	[(None, 32, 32, 3)]	0
conv2d (Conv2D)	(None, 32, 32, 64)	1792
conv2d_1 (Conv2D)	(None, 32, 32, 64)	36928
max_pooling2d (MaxPooling2D)	(None, 16, 16, 64)	0
conv2d_2 (Conv2D)	(None, 16, 16, 128)	73856
conv2d_3 (Conv2D)	(None, 16, 16, 128)	147584
max_pooling2d_1 (MaxPooling2D)	(None, 8, 8, 128)	0
conv2d_4 (Conv2D)	(None, 8, 8, 256)	295168
conv2d_5 (Conv2D)	(None, 8, 8, 256)	590080
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 256)	0
conv2d_6 (Conv2D)	(None, 4, 4, 512)	1180160
conv2d_7 (Conv2D)	(None, 4, 4, 512)	2359808
dropout (Dropout)	(None, 4, 4, 512)	0

max_pooling2d_3 (MaxPooling 2D)	(None, 2, 2, 512)	0
conv2d_8 (Conv2D)	(None, 2, 2, 1024)	4719616
conv2d_9 (Conv2D)	(None, 2, 2, 1024)	9438208
dropout_1 (Dropout)	(None, 2, 2, 1024)	0
flatten (Flatten)	(None, 4096)	0
dense (Dense)	(None, 512)	2097664
dense_1 (Dense)	(None, 256)	131328
dense_2 (Dense)	(None, 1)	257
<hr/>		
Total params:	21,072,449	
Trainable params:	21,072,449	
Non-trainable params:	0	

Training:

Split the dataset into training and validation sets for model training and evaluation, respectively. Train the U-Net model using the training dataset. Monitor the loss and validation accuracy during training to ensure the model is learning effectively. Experiment with different hyperparameters, such as learning rate, batch size, and number of epochs, to optimize the model performance. Since the dataset is imbalanced, consider using techniques like class weighting or oversampling to handle class imbalance during training. Evaluation:

```
# Calculate steps per epoch
batch_size= 32
steps_per_epoch = len(train_images) // batch_size
# Train the model
history = model.fit(
    train_generator,
    steps_per_epoch=steps_per_epoch,
    epochs=10,
    validation_data=(test_images, test_labels),
    class_weight=class_weights
)
→ Epoch 1/10
2024-03-19 23:55:48.944144: I tensorflow/core/common_runtime/executor.cc:1197] [/device:CPU:0] (DEBUG INFO) Executor start a
[[{{node Placeholder/_0}}]]
2/2 [=====] - 3s 590ms/step - loss: 0.5143 - accuracy: 0.3488 - val_loss: 9.6630 - val_accuracy: 0.
Epoch 2/10
2/2 [=====] - 1s 341ms/step - loss: 0.3705 - accuracy: 0.9535 - val_loss: 1.1406 - val_accuracy: 0.
Epoch 3/10
2/2 [=====] - 1s 333ms/step - loss: 0.0984 - accuracy: 0.9302 - val_loss: 0.4167 - val_accuracy: 0.
Epoch 4/10
2/2 [=====] - 2s 776ms/step - loss: 0.1243 - accuracy: 0.9375 - val_loss: 0.4816 - val_accuracy: 0.
Epoch 5/10
2/2 [=====] - 1s 773ms/step - loss: 0.0562 - accuracy: 0.9070 - val_loss: 1.1776 - val_accuracy: 0.
Epoch 6/10
2/2 [=====] - 1s 356ms/step - loss: 0.0483 - accuracy: 0.9535 - val_loss: 1.4093 - val_accuracy: 0.
Epoch 7/10
2/2 [=====] - 1s 368ms/step - loss: 5.0818e-08 - accuracy: 1.0000 - val_loss: 1.3731 - val_accuracy
Epoch 8/10
2/2 [=====] - 1s 340ms/step - loss: 0.0739 - accuracy: 0.9302 - val_loss: 0.9842 - val_accuracy: 0.
Epoch 9/10
2/2 [=====] - 2s 772ms/step - loss: 0.0428 - accuracy: 0.9375 - val_loss: 0.6000 - val_accuracy: 0.
Epoch 10/10
2/2 [=====] - 2s 792ms/step - loss: 0.0329 - accuracy: 0.9219 - val_loss: 0.4146 - val_accuracy: 0.
```

Evaluate the trained model on the holdout dataset to assess its performance. Calculate metrics such as precision, recall, F1-score, and PR-AUC to measure the model's effectiveness in detecting TB bacilli. Visualize the model predictions on sample images to understand its behavior qualitatively.

```
from sklearn.metrics import roc_curve, roc_auc_score
import matplotlib.pyplot as plt

# 1. Evaluate the model on the test dataset
predictions = model.predict(test_images)

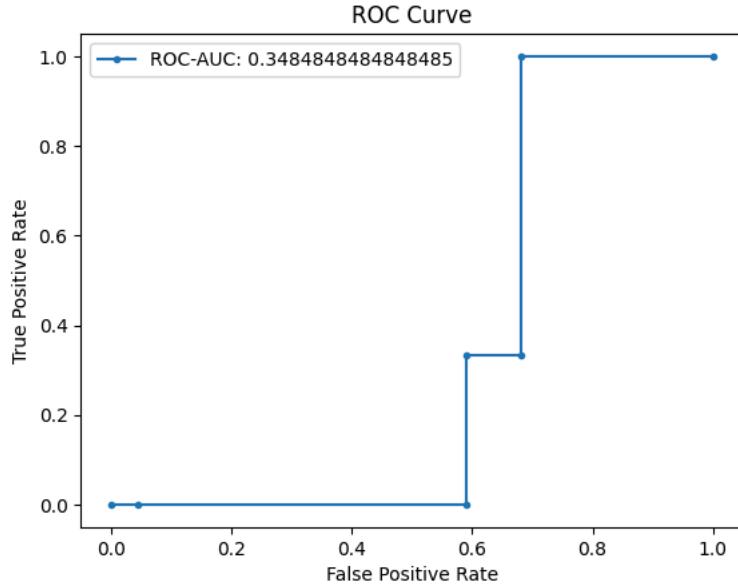
# 1. Calculate ROC curve
fpr, tpr, _ = roc_curve(test_labels, predictions)

# 2. Calculate ROC-AUC score
roc_auc = roc_auc_score(test_labels, predictions)

# 3. Plot ROC curve
plt.plot(fpr, tpr, marker='.', label=f"ROC-AUC: {roc_auc}")
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('ROC Curve')
plt.legend()
plt.show()

print(f"ROC-AUC: {roc_auc}")
```

→ 1/1 [=====] - 0s 163ms/step



ROC-AUC: 0.3484848484848485

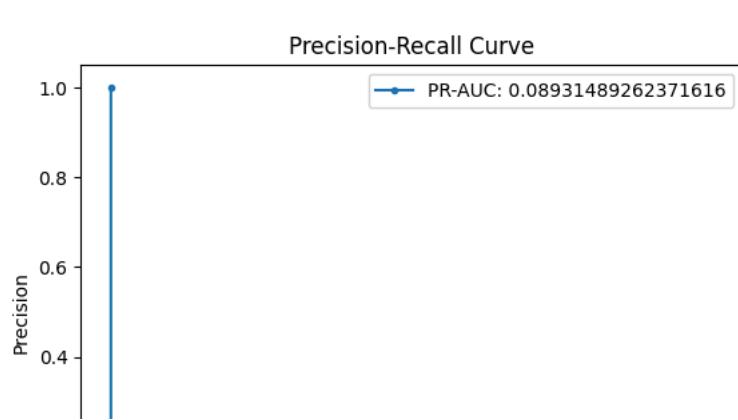
```
from sklearn.metrics import precision_recall_curve, auc

# 1. Calculate precision-recall curve
precision, recall, _ = precision_recall_curve(test_labels, predictions)

# 2. Calculate PR-AUC score
pr_auc = auc(recall, precision)

# 3. Plot PR curve
plt.plot(recall, precision, marker='.', label = f"PR-AUC: {pr_auc}")
plt.xlabel('Recall')
plt.ylabel('Precision')
plt.title('Precision-Recall Curve')
plt.legend()
plt.show()

print(f"PR-AUC: {pr_auc}")
```



Scoring Submission Prediction

```
csv_file = '/home/ngsci/datasets/tb-wellgen-smear/supplementary/contest/tb-holdout-manifest.csv'  
holdout_df = pd.read_csv(csv_file)  
holdout_df=holdout_df[:100]  
holdout_df
```

	image_id	file_path
0	0002da84	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
1	0002e02e	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
2	00042206	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
3	00072c47	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
4	0008332c	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
...
95	010f0313	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
96	011336c4	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
97	011480fe	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
98	0117448f	/home/ngsci/datasets/tb-wellgen-smear/suppleme...
99	011889dc	/home/ngsci/datasets/tb-wellgen-smear/suppleme...

100 rows × 2 columns

```
# Preprocess images for holdout without augmentation  
holdout_images = np.array([preprocess_image(file_path) for file_path in holdout_df['file_path']])
```

```
# 1. Evaluate the model on the holdout dataset  
predictions = model.predict(holdout_images)
```

```
4/4 [=====] - 0s 23ms/step
```

```
holdout_df['probability'] = predictions
```

```
# Drop the original file_path column  
holdout_df.drop(columns=['file_path'], inplace=True)
```

```
# Save the modified holdout DataFrame to CSV  
modified_holdout_csv_path = "modified_holdout.csv"  
holdout_df.to_csv(modified_holdout_csv_path, index=False)
```