

Article

A Cytopathologist Eye Assistant for Cell Screening

Débora N. Diniz ^{1,*}, Breno N. S. Keller ^{1,†}, Mariana T. Rezende ², Andrea G. C. Bianchi ¹,
Claudia M. Carneiro ², Renata R. e R. Oliveira ², Eduardo J. S. Luz ¹, Daniela M. Ushizima ^{3,4,5},
Fátima N. S. de Medeiros ⁶ and Marcone J. F. Souza ¹

- ¹ Departamento de Computação, Universidade Federal de Ouro Preto (UFOP), Ouro Preto 35400-000, Brazil
² Departamento de Análises Clínicas, Universidade Federal de Ouro Preto (UFOP), Ouro Preto 35400-000, Brazil
³ Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
⁴ Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA
⁵ Bakar Computational Health Sciences Institute, University of California, San Francisco, CA 94143, USA
⁶ Departamento de Engenharia de Teleinformática, Universidade Federal do Ceará (UFC), Fortaleza 60455-970, Brazil
* Correspondence: debora.diniz@aluno.ufop.edu.br
† These authors contributed equally to this work.



Citation: Diniz, D.N.; Keller, B.N.S.; Rezende, M.T.; Bianchi, A.G.C.; Carneiro, C.M.; Oliveira, R.R.e.R.; Luz, E.J.S.; Ushizima, D.M.; de Medeiros, F.N.S.; Souza, M.J.F. A Cytopathologist Eye Assistant for Cell Screening. *AppliedMath* **2022**, *2*, 659–674. <https://doi.org/10.3390/appliedmath2040038>

Academic Editor: Takayuki Hibi

Received: 1 November 2022

Accepted: 25 November 2022

Published: 30 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Screening of Pap smear images continues to depend upon cytopathologists' manual scrutiny, and the results are highly influenced by professional experience, leading to varying degrees of cell classification inaccuracies. In order to improve the quality of the Pap smear results, several efforts have been made to create software to automate and standardize the processing of medical images. In this work, we developed the CEA (Cytopathologist Eye Assistant), an easy-to-use tool to aid cytopathologists in performing their daily activities. In addition, the tool was tested by a group of cytopathologists, whose feedback indicates that CEA could be a valuable tool to be integrated into Pap smear image analysis routines. For the construction of the tool, we evaluate different YOLO configurations and classification approaches. The best combination of algorithms uses YOLOv5s as a detection algorithm and an ensemble of EfficientNets as a classification algorithm. This configuration achieved 0.726 precision, 0.906 recall, and 0.805 F1-score when considering individual cells. We also made an analysis to classify the image as a whole, in which case, the best configuration was the YOLOv5s to perform the detection and classification tasks, and it achieved 0.975 precision, 0.992 recall, 0.970 accuracy, and 0.983 F1-score.

Keywords: cancer cell detection; Pap smear image; cervical cytology; deep learning; decision support tool; whole image analysis

1. Introduction

Pap smear has been valuable for the prevention of cervical cancer for decades [1]. Unfortunately, this neoplasm is still one of the main causes of death from cancer among the female population in developing countries [2]. Manual Pap smear analysis presents several issues, such as the number of false-negative results of up to 62%, which is a serious problem because it can delay a treatable early detection case [3–9].

One of the leading causes is the visual inspection of cytopathological lesions under an optical microscope, which comprises two main tasks performed by the cytopathologist: locating and classifying pre-neoplastic cells. At this examination stage, scrutiny and interpretation errors may occur, respectively [10].

Excellence in the performance of these tasks is closely related to the professional's qualification and experience and can also be hampered by the extensive daily workload, fatigue, qualitative cytomorphological criteria, and the wide variety of morphological lesions in different samples [11]. Therefore, cytopathological diagnosis is linked to subjective factors that culminate in human errors [12].

Thus, monitoring the quality of the Pap smear has allowed the application of a set of measures aimed at reducing deficiencies in the process [13]. Computer-aided diagnosis (CAD) has gained evidence in this context as these systems can achieve promising results in minimizing diagnostic errors [14], acting as a supplementary tool for decision-making.

Computational approaches in this context need to deal with a different set of challenges. The main one is the overlapping elements within images from conventional cytology, which makes it challenging to perform classification and detection algorithms. This problem can be circumvented by using images collected in liquid-based cytology since they have fewer overlapping elements than in conventional cytology. However, this approach for generating images has high costs and is not widely used.

This work proposes a tool to support the cytopathologist's decision through automated analysis of images based on deep learning, which provides feedback to the user about the possible classifications of cervical cells from the Pap smear slide. The cytopathologist uses the CAD system output as a "second opinion" and makes the final decision. Our contributions are summarized as follows:

- (A) Detection and classification of cells in multi-cell images that do not require prior segmentation and hand-crafted features;
- (B) Evaluation of different YOLO architectures and their configurations to detect and classify the nuclei from the CRIC Cervix images;
- (C) Introduction of a computer-aided diagnosis system to detect and classify cervical cells, the Cytopathologist Eye Assistant (CEA).

The methodology proposed in this work is limited to databases with known ground truth nuclei. Furthermore, supervised machine learning models need to receive data similar to those they have been trained on; otherwise, they may perform poorly. So, in this case, for the models presented to continue with good results, the images must have the same characteristics, such as but not limited to resolution, zoom, and focus. Ideally, they should be collected on the same equipment used to construct the databases.

The remainder of this paper is organized as follows. Section 2 presents the related studies, followed by the details of the proposed methodology in Section 3. Section 4 shows our experiments, including the database, metrics, results, and discussions. Section 5 introduces the CEA tool, our proposal for a computer-aided cervical cell screening system. Finally, the conclusions are exposed in Section 6.

2. Background

This section is organized as follows. Section 2.1 reviews methods of cell detection and classification, while Section 2.2 presents an analysis of computational tools to support the cytopathologist's work routine. Finally, Section 2.3 presents a review of YOLOv5 architecture.

2.1. Detection and Classification Review

Recently, several studies have analyzed different approaches to support cytopathologists in their work routine. One of these approaches is automatic cell classification since this task is an essential part of the professional's routine and is also challenging. Rehman et al. [15] proposed a method to extract deep features in cells and used softmax regression, support vector machine, and GentleBoost ensemble of decision trees to perform the classification. Ghoneim et al. [16] also performed feature extraction for classification using Shallow, VGG-16, and CaffeNet architectures and they used Extreme Learning Machine and Autoencoder for classification. Diniz et al. [17] investigated ten deep convolutional neural networks and proposed an ensemble of the three best ones to perform the cell classification.

Some studies performed cell classification based on handcrafted features. Diniz et al. [18] investigated eight traditional machine learning methods to perform hierarchical cell classification. Di Ruberto et al. [19] used different descriptors to classify the images using k-nearest neighbors, and Bora et al. [20] classified cells based on their shape, texture, and color using an ensemble method.

However, performing the automatic classification of cervical lesions is a highly challenging task that becomes even more challenging when combined with automatic cell detection.

Matias et al. [21] presented a survey that addresses cytology assisted by computer vision in a whole slide image. They reported that all studies employed state-of-the-art object detectors instead of proposing a new architecture of Convolutional Neural Network (CNN). An example of this approach is presented by Xiang et al. [22], who used YOLOv3 followed by an additional task-specific classifier to perform the detection and classification.

The studies presented in this review use unicellular databases and/or private databases, and/or liquid-based cytology databases. Single-cell databases have the disadvantage of not representing reality since unit cells do not overlap. Private databases do not allow a good comparison for future work. Moreover, liquid-based cytology databases have images that are easier to automatically classify since the collected cells undergo a preparation to improve their spatial distribution. However, this preparation increases the exam cost, making it less used in middle and low-income countries.

Unlike the studies presented, we investigated convolutional neural networks to perform image detection and classification using a newly published multi-cell image database of conventional cytology, the most used in underdeveloped countries. Furthermore, we used the cell nuclei as the region of interest for the detection and classification tasks since Diniz et al. [18] showed that this region contains enough information to perform the classification process. In addition, we presented a decision support tool to complement the work, followed by feedback from cytopathologists who evaluated the tool.

2.2. Landscape Analysis

Despite many scientific publications on cervical cancer and artificial intelligence (AI) applied to cell lesion detection, designing and adopting diagnostic software is still a major challenge. Two medical companies stand out with the following software, BD Focal Point Slide Profiler (<https://www.bd.com/en-us/products-and-solutions/products/product-families/bd-focalpoint-gs-imaging-system>) (accessed on 1 November 2022) and Thin Prep Imaging System (TIS) (<http://www.hologic.ca/products/clinical-diagnostics-and-blood-screening/instrument-systems/thinprep-imaging-system>) (accessed on 1 November 2022), both on the market since 1980.

The BD FocalPoint Slide profiler is attached to a slide scanner [23]. The slides are analyzed, and the software returns information about the likelihood that the slide contains abnormalities. This information is sent to the professional for review. The system returns regions of interest to the user instead of the entire slide, facilitating the process and reducing the workload.

The TIS system combines image processing and slide scanning automation [24]. It allows the analysis of 22 regions of interest and forwards the results to the professional who can intuitively track them. Each slide is sent to the professional for review, and complex areas can be marked for further analysis.

Recently, Hologic introduced GeniusTM Digital Diagnostics System, a digital cytology platform that combines a new artificial intelligence algorithm with advanced volumetric imaging technology to help pathologists identify pre-cancerous and cancerous lesions in women [25]. The AI algorithm identifies suspicious structures on the slide. The regions of interest are classified, and the most relevant lesions are presented to the pathologist for review. For now, Genius diagnostic software is only being sold in Europe.

We also highlight the company KFBIO (<https://www.kfbiopathology.com/>) (accessed on 1 November 2022), producer of the KFBIO Pathological Remote Diagnosis Platform software that works with cervical cancer and is coupled to a slide scanner, allowing high efficiency, sensitivity of 98.2%, and specificity of 63.5%. Artificial intelligence is applied based on deep convolutional neural networks to analyze liquid-based cervical cytology samples to confirm their feasibility in clinical practice.

All approaches are linked to the whole slide imaging (WSI) scanning equipment, a complete microscope slide scanning, and a high-resolution digital image. Thus, their use

is restricted to the acquisition of high-cost equipment. In addition, their use is limited to liquid-based slides and Thin Prep tests, so it is impossible to use samples obtained by the conventional method.

In 2019, a startup (<https://datexim.ai/>) (accessed on 1 November 2022) presented the CytoProcessor™ software that uses artificial intelligence methods to differentiate normal cells from lesioned cells and then performs classification [26]. Cells are shown in a table sorted from the most to the least diagnostically relevant. Each cell can be visualized in its context, similar to the microscope view. Compared to TIS, CytoProcessor™ significantly improves diagnostic sensitivity without compromising specificity [27].

Some computational tools even analyze cell structures individually or in collections, which may be helpful during the diagnostic process. However, none are specific to support the diagnosis of cervical cancer.

QuPath (<https://qupath.github.io/>) (accessed on 1 November 2022), Orbit (<http://www.orbit.bio/>) (accessed on 1 November 2022), and SlideRunner (<https://github.com/DeepPathology/SlideRunner>) (accessed on 1 November 2022) softwares are open-source and free. They allow manual and automatic markings on biological images obtained by WSI. They also support analysis of AI algorithms, but its use is not trivial and requires professional programming.

Another free and open-source software is Cytomine (<https://cytomine.com/>) (accessed on 1 November 2022), a web platform that promotes collaborative image analysis through segmentation, manual cell annotation services and collection generation. In addition, it offers paid AI analysis services for images.

As mentioned, the software allows integration with image analysis algorithms. However, such use must be implemented separately and integrated into the platforms, which is not trivial. None of them have AI tools that allow support for diagnosing a cervical cell lesion quickly and easily for the pathologist without the support of an AI specialist.

2.3. YOLOv5 Review

The network architecture of YOLOv5 is composed of three components: backbone, neck, and head. The backbone component processes the image to generate feature sets at different levels. In YOLOv5, the backbone component is made by CSPDarknet [28] architecture with extracted features based on an adapted DenseNet [29] process. Its architecture is shown in Table 1.

Table 1. YOLOv5 backbone architecture.

Stage	Operator	#Channels	#Layers
1	Focus	64	1
2	Conv	128	1
3	BottleneckCSP	128	3
4	Conv	256	1
5	BottleneckCSP	256	9
6	Conv	512	1
7	BottleneckCSP	512	9
8	Conv	1024	1
9	SPP	1024	1
10	BottleneckCSP	1024	3

The neck component combines the features generated in the backbone and passes them to the prediction task. Therefore, this component uses the PANet [30] architecture to create feature pyramids, which can represent objects in different scales and work with images of different sizes. In addition, the use of PANet improves the information flow.

Finally, the head component, performed by the YOLO layer, consists of using the features extracted by the previous components to generate the bounding box, the probability of each class, and the anchors of each detected object in the image. The hidden layers on the network and the detection layer use as an activation function, the ReLU function, and the Sigmoid function, respectively.

2.4. YOLOR Review

The YOLOR architecture is adapted from the YOLOv4 and works as an architecture to perform multiple tasks in a single execution. It combines the YOLOv4 architecture with the concepts of implicit and explicit knowledge in its execution [31]. An overview of its architecture is presented in Figure 1.

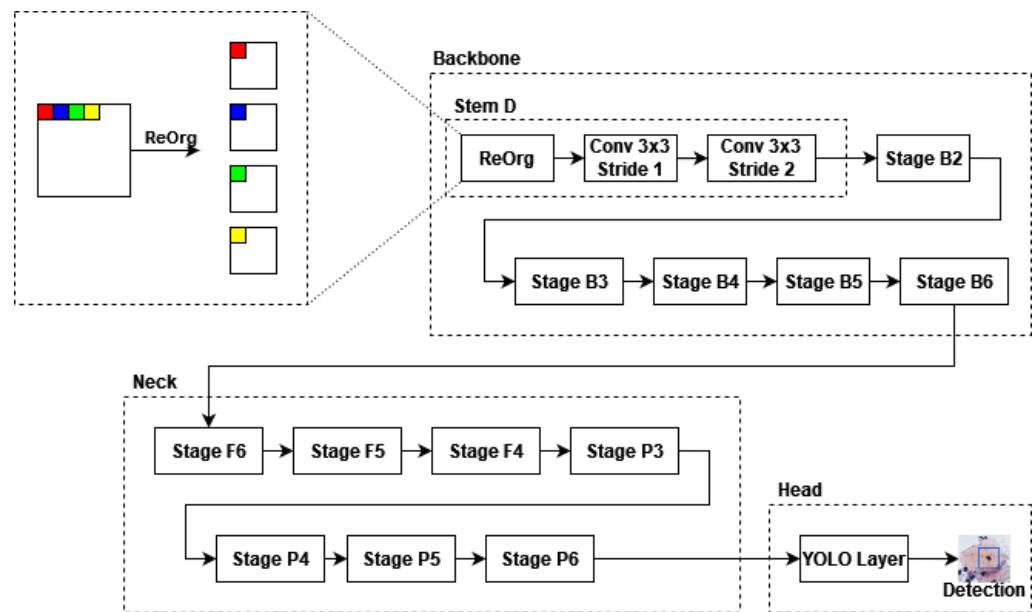


Figure 1. YOLOR architecture, adapted from Wang et al. [31].

According to Wang et al. [31]: (i) the backbone component is made by a focus layer (Stem D) followed by Stages B2, B3, B4, B5, and B6; (ii) YOLOR uses a CSP convolution as a down-sampling module in each stage to optimize the gradient propagation; (iii) the base channels and repeat number of the backbone layers are set to $\{128, 256, 384, 512, 640\}$ and $\{3, 7, 7, 3, 3\}$, respectively; and, (iv) one difference between YOLOR and YOLOv4 is that YOLOv4 uses Mish as an activation function, while YOLOR uses SiLU.

3. Methodology

Our proposal applies a CNN to detect and classify regions of interest and presents the results to assist cytopathologists in interpreting cervical cells. We used neural networks based on the YOLO (You Only Look Once) [32,33] architecture, which presents excellent results and shorter times for training and inference. Furthermore, as YOLO considers the image as a whole, this also allows it to reduce the occurrence of background errors compared to other detection networks such as R-CNN and Fast R-CNN [32].

This work uses YOLOv5 [34] and YOLOR [31] implementations, which present better inference times than other detection and classification networks [35]. In addition, they present a more modularized way to run tests and detections using the architecture. These characteristics were crucial for choosing these networks as they allow easier integration into a tool.

This modularization of the YOLO architecture also allows using other classification algorithms to classify the detected objects. Based on that, this work also evaluated the

combination of the YOLO detection and the classification proposed by Diniz et al. [17], which obtained the best results for the database used.

Figure 2 presents the general workflow of this proposal, which consists of three steps: (1) Model training and evaluation; (2) Encapsulation of the model within a software interface; and (3) Assessment of the tool and model results with specialists. In Step (1), we use the CRIC Cervix to train and evaluate base models previously selected. We show these models in Section 4. This step is associated with our contributions A, B, and C. In Step (2), a tool named CEA was built with the best performance configuration obtained in Step (1) and is intended to offer the professionals the model results in a friendly and accessible interface. This step is associated with our contributions D and E. Finally, in Step (3), specialists evaluated the tool and provided feedback on its use.

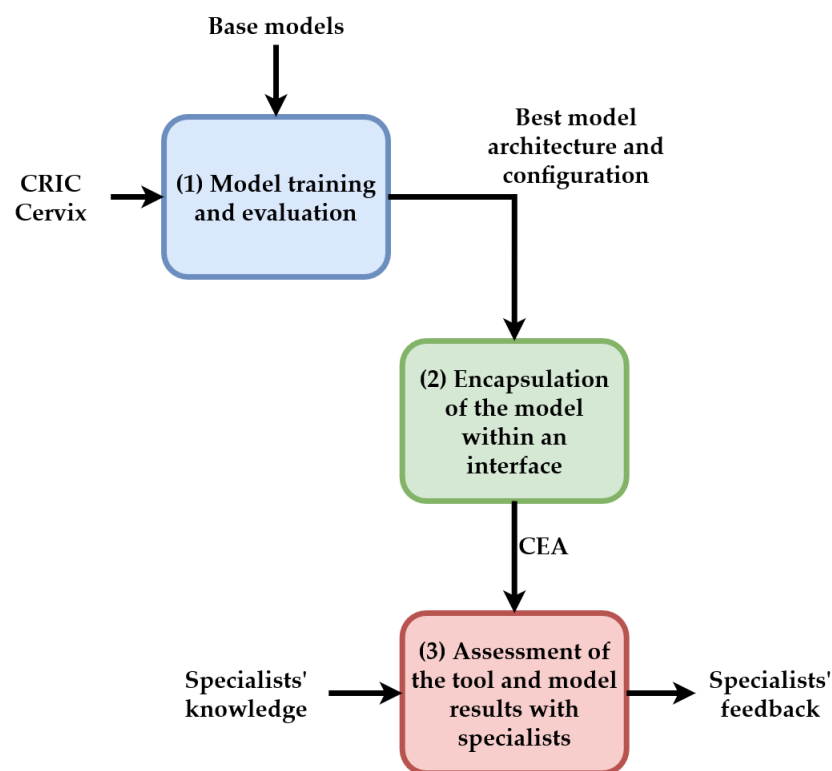


Figure 2. Our proposal workflow.

4. Experiments

This section describes how our experiments were constructed and the results achieved.

4.1. Database

We tested the proposed method on the cervical cell classification collection available in the CRIC Searchable Image Database (<https://database.cric.com.br>) (accessed on 1 November 2022), the CRIC Cervix [36]. This collection contains 400 images from slides of the conventional smear (Pap smear) captured using conventional microscopy in a bright field with a 40× objective and a 10× eyepiece through a Zeiss AxioCam MRc digital camera coupled to a Zeiss AxioImager Z2 microscope with the Axio Vision Zeiss Software.

The collection images have a totality of 11,534 marks of their cells represented by bounding boxes of 100 px × 100 px. Four cytopathologists classified and revised each mark according to the Bethesda System nomenclature for cell lesions, labeled as Negative for Intraepithelial Lesion or Malignancy (NILM), Atypical Squamous Cells of Undetermined Significance (ASC-US), Low-grade Squamous Intraepithelial Lesion (LSIL), Atypical Squamous Cells, cannot exclude HSIL (ASC-H), High-grade Squamous Intraepithelial Lesion (HSIL), and Squamous Carcinoma (SCC). NILM labels correspond to normal cells, and the others (ASC-US, LSIL, ASC-H, HSIL, and SCC) correspond to lesioned cells (distinct

grades). Figure 3 shows an image sample from the cervical cell classification collection, where colored boxes represent different lesions. A more detailed database description may be found in [36].

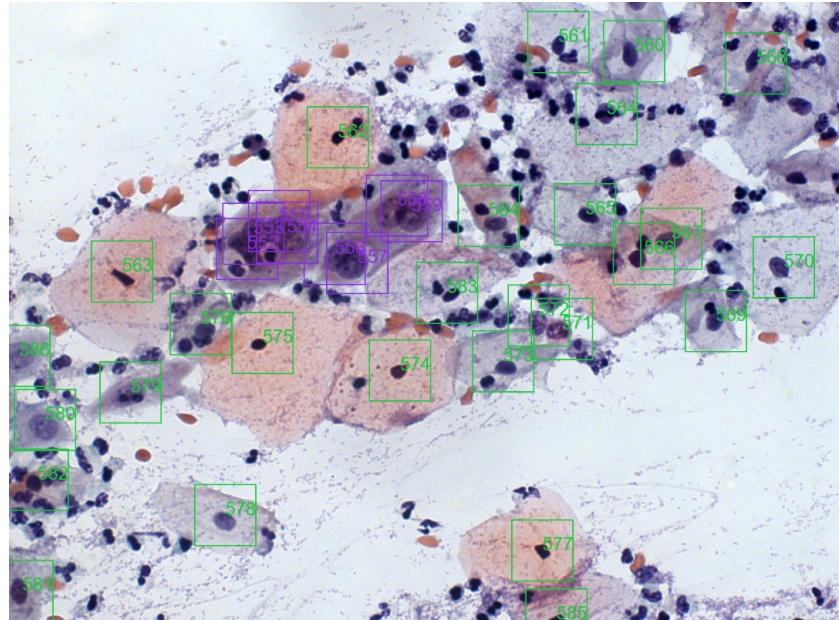


Figure 3. Image example from the cervical cell classification collection.

4.2. Database Manipulations

For the image's configuration, we considered: the original format of the images, resizing the images to 561 px × 451 px, and the generation of new images for training using augmentation strategies and balancing classes.

We utilized an undersampling method for balancing the database. That method removed images that contained many samples of the negative class to balance the number of examples for training.

The augmentation was implemented using Python's Clodsa library [37]. We used the horizontal, vertical, and vertical–horizontal flips, average blurring (*kernel* parameter equals 5), and raised hue (*power* parameter as 0.9).

With these image manipulations, the new database with the resized images had the same size as the original database, with 400 images. Simultaneously, the database with resized images and augmentation comprised 1800 images, and the database with resized images and balanced classes comprised 332 images.

4.3. Metrics

We apply precision, recall, and mAP (mean average precision) metrics to evaluate the proposed solution using CNN. Precision (see Equation (1)) measures the model's ability to detect altered cells only in lesioned slides; recall (see Equation (2)) measures the model's ability to detect altered cells; mAP (see Equation (3)) measures the model's performance in detecting and retrieving information in the images. These metrics are typically used in object detection and classification scenarios.

$$Precision = \frac{|TP|}{|TP| + |FP|} \quad (1)$$

$$Recall = \frac{|TP|}{|TP| + |FN|} \quad (2)$$

$$mAP = \frac{1}{|Classes|} \sum_{c \in Classes} \frac{|TP_c|}{|TP_c| + |FP_c|} \quad (3)$$

In these equations, $|TP|$, $|FP|$, $|FN|$, and $|Classes|$ represent the number of true positives, false positives, false negatives, and the number of classes, respectively.

4.4. Development and Test Environment

We used a computer with an Intel Core i7-9700F processor with a GPU GeForce RTX 2080 and 16 GB of RAM in a Windows 64-bit system to perform the experiments. Algorithms were written in Python and leveraged the PyTorch framework, versions 3.7.9 and 1.7, respectively.

We split the database into training, testing, and validation sets. The distribution of each database setup is shown in Table 2. We also performed a 10-fold evaluation as suggested by Dhurandhar et al. [38]. All results are the average obtained by cross-validation.

Table 2. Database setup distribution.

Database Setup	Train	Test	Validation
Original	280	40	80
Resized	280	40	80
Resized and Augmentation	1680	40	80
Resized and Balanced	233	33	66

4.5. Configuration and Evaluation of the YOLO Models

In the first four columns of Table 3, we present the combinations between the base model, the image configuration, and the model's fitness formula used in our experiments. The models considered were the small (*s*), medium (*m*), large (*l*), and extra-large (*x*) YOLOv5 configurations and the YOLOR-P6 model.

Table 3. Results of the object detection and classification.

Model	Database Setup	Fitness	Batch	Precision	Recall	mAP@0.5	F1-Score
YOLOv5s	Original Database	Default	8	0.736	0.777	0.806	0.756
YOLOv5l	Resized and Augmentation	Default	16	0.767	0.736	0.799	0.751
YOLOv5m	Resized and Augmentation	Default	16	0.748	0.734	0.799	0.741
YOLOv5x	Resized and Augmentation	Default	8	0.760	0.718	0.801	0.738
YOLOv5s	Resized and Augmentation	Default	16	0.734	0.727	0.767	0.730
YOLOv5s	Original Database	Equal	8	0.743	0.718	0.780	0.730
YOLOv5s	Original Database	Recall	8	0.719	0.727	0.780	0.723
YOLOv5x	Resized	Default	8	0.719	0.702	0.742	0.710
YOLOR	Resized	Default	8	0.536	0.859	0.784	0.660
YOLOR	Resized and Balanced	Default	8	0.499	0.882	0.792	0.637
YOLOR	Resized and Balanced	Recall	8	0.223	0.884	0.671	0.356
YOLOR	Resized	Recall	8	0.146	0.912	0.673	0.252

We use pre-trained networks, which means that the networks used have been previously trained on a large dataset. The advantage of this pre-training is the improvement of the model's capability to extract basic features. This happened because, in CNNs, the features are more generic in the early layers and more specific to the treated problem in the last layers [39]. Therefore, it is possible to fine-tune a generic network on a smaller dataset of the target domain using transfer learning. So, the YOLOv5 was pre-trained with the ImageNet [40] database and the YOLOR with MS COCO [41] database.

For the fitness formula, we considered: the default formula (“Default” in Table 3) that combines mAP@0.5 and mAP@0.5:0.95; a variation that considers equal weights for all metrics (“Equal” in Table 3), therefore no priority in metrics; and a variation prioritizing the recall (“Recall” in Table 3), since this metric is the most important for the problem addressed as it measures the amount of lesioned cells that were classified as normal.

The networks were trained with the following hyperparameters: learning rate of 0.01, momentum of 0.937, weight_decay of 0.0005, and IoU threshold (in training) of 0.2 and considering rectangular bounding boxes. An IoU threshold of 0.45 and a confidence threshold of 0.25 were considered for the evaluation. Values for these parameters are from the default network configuration.

In addition, the highest possible value was used for the batch, which is limited by the model size in the GPU memory. We also emphasize that not all combinations could be tested due to memory constraints.

The last four columns of Table 3 summarize our results ordered by the F1-score metric. The best result for each metric is highlighted in bold. Each model was trained for 200 epochs. The best result, considering the F1-score metric for the evaluated combinations (see Table 3), was obtained by the YOLOv5s model with the original image and the model’s default fitness. Under these conditions, the recall was 0.777 for the test set. However, the best recall value obtained among the evaluated combinations was 0.884. This result was achieved through the YOLOR model using recall as the only metric in the model’s fitness. In contrast, when performing this prioritization, there was a sharp loss of precision.

The same observation holds for precision. The model with the best F1-score metric has a precision of 0.736, while the best precision found was 0.767. However, this increase in accuracy resulted in a loss in the recall. When considering the F1-score, the result of the YOLOv5s model was better than those obtained by the YOLOv5l model. As previously stated, recall is more important for the considered problem, mainly because, in an actual scenario, the tool detections would go through the validation of a professional. The false positives that influenced the accuracy value would be easily excluded by the professional.

It is also important to emphasize that the resized image allows larger models (due to memory allocation). However, the use of these models did not lead to better results. Looking at Table 3, we can see that the configurations *m*, *l*, and *x* of the YOLOv5 model were better than those of the configuration *s* for the resized images. However, the best result was obtained with the images in the original size. Thus, the benefit of resizing the images is restricted to reducing the model training time and the necessary memory allocation.

4.6. Comparison between Different Classification Algorithms

Considering only the nuclei classification of the CRIC Database images, we found that in the literature, the best result was achieved by Diniz et al. [17]. In their work, the authors proposed an ensemble of EfficientNets to perform the classification. Thus, we analyzed if using this ensemble as a classifier of the objects detected by YOLO improves its results. We also investigate each EfficientNet that composes the ensemble method as a classifier. Table 4 reports the obtained results of this investigation. As Table 3, the results are ordered by the F1-score metric, and the best ones are highlighted in bold.

As we can see in Table 4, performing nuclei detection using YOLOv5s and the classification using the ensemble proposed by Diniz et al. [17], there was an improvement in recall (from 0.777 to 0.906) and F1-score (from 0.756 to 0.805). Precision dropped from 0.736 to 0.726, and AUC dropped from 0.825 to 0.816.

Table 4. Results of the object's classification.

Detection Model	Classification Model	Precision	Recall	F1-Score	AUC
YOLOv5s	Ensemble [17]	0.726	0.906	0.805	0.816
YOLOv5s	EfficientNetB6 [17]	0.745	0.889	0.804	0.821
YOLOv5s	EfficientNetB2 [17]	0.720	0.898	0.797	0.808
YOLOv5s	EfficientNetB1 [17]	0.698	0.896	0.782	0.792
YOLOv5s	YOLOv5s	0.736	0.777	0.756	0.825

4.7. Whole Image Classification

We also evaluate the whole image classification; that is, if at least one nucleus of the image is classified as lesioned, the image is also classified as lesioned. On the other hand, if all nuclei are classified as normal, the image is also considered normal. This evaluation corresponds to an image classification as a whole, similar to how professionals do their daily work. So, we performed and evaluated the whole image classification for the best approaches presented in Sections 4.5 and 4.6: the YOLOv5s performing the classification and detection of the nuclei; and the combination of the YOLOv5s for the detection and the EfficientNets ensemble for the classification.

Table 5 presents the results (ordered by F1-score with the best ones highlighted in bold) for this general image evaluation. Using YOLOv5s for detecting and classifying nuclei, it achieved the best precision and F1-score results. Its recall was lower than the other approach, but it remained high. For the AUC metric, both scenarios obtained unsatisfactory results. However, this may have been caused by the distribution of samples of each class in the test set. When we modify the problem to classify the image as a whole, we have very few samples of entirely negative images. This limitation is due to the database that was planned for classifying lesioned nuclei in an image since images containing only non-lesioned nuclei are underinvestigated.

Table 5. Results of the whole image classification.

Detection Model	Classification Model	Precision (σ)	Recall (σ)	F1-Score (σ)	AUC (σ)
YOLOv5s	YOLOv5s	0.975 (0.021)	0.992 (0.013)	0.983 (0.011)	0.493 (0.012)
YOLOv5s	Ensemble [17]	0.878 (0.028)	1.000 (0.000)	0.934 (0.016)	0.500 (0.000)

4.8. Literature Comparison

Carrying out a comparison with the results obtained by Diniz et al. [17] is not feasible because the authors only dealt with the classification task, and the results obtained in our work carry the error from both detection and classification tasks. When writing this work, we did not find any work in the literature that detected and classified nuclei in the CRIC Cervix database.

Xiang et al. [22] presented a similar approach for this situation, obtaining an mAP of 0.634 and a recall of 0.975 in the whole image analysis. Our results of mAP of 0.806 and recall of 0.992 were superior to theirs, but it is important to note that the bases used in these two studies differ. It is worth emphasizing that Xiang et al. [22] used liquid-based cytology images, which means that the collected cells were pre-processed to improve their spatial distribution leading to an easier problem than the one we approach in this work.

5. CEA: Cytopathologist Eye Assistant

We developed the Cytopathologist Eye Assistant (CEA) using Python's Kivy library (<https://kivy.org>) (accessed on 1 November 2022). The idea of the tool is to assist the cytopathologist in diagnosing cervical cancer using images obtained from the Pap smear. In general, the CEA follows the interaction displayed in Figure 4, in which a specialist

inputs an image for the system, and a detection algorithm processes this image. The results of the detection are the input for a classification algorithm. In the end, both results are combined and shown for the specialist in the CEA interface. In addition, the detection and classification models can be changed without impacting the general interaction flow of the tool; however, the models need to accept an image as input and provide bounding boxes and labels as output.

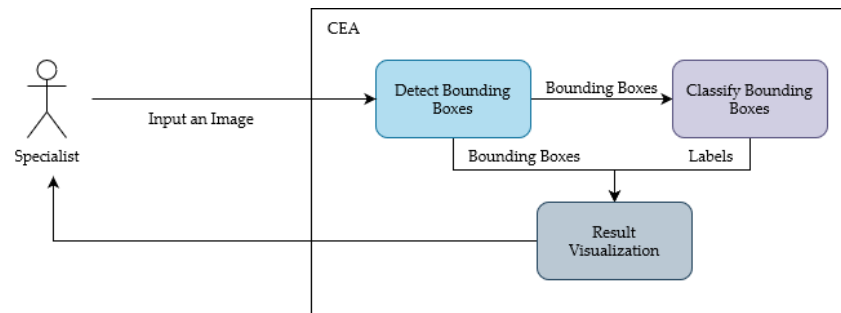


Figure 4. CEA interaction flow.

For implementation, we have used the YOLO model with the best result configuration obtained in this work. This model was selected because YOLO and the ensemble approach have similar results for the whole image classification (Table 5). However, YOLO has the detection and classification unified in a single process improving the tool processing time.

The CEA interface is presented in Figures 5 and 6. Its functionalities are highlighted through numbered red circles. In Figure 5, the user clicks on the “Upload your image” button to select the image to have its cells detected and classified (see red circle 1). CEA shows the chosen image below (see red circle 2), and the user can send the selected image for processing (by clicking on the “Process your image” button, see red circle 3) or can select another one (in red circle 1).

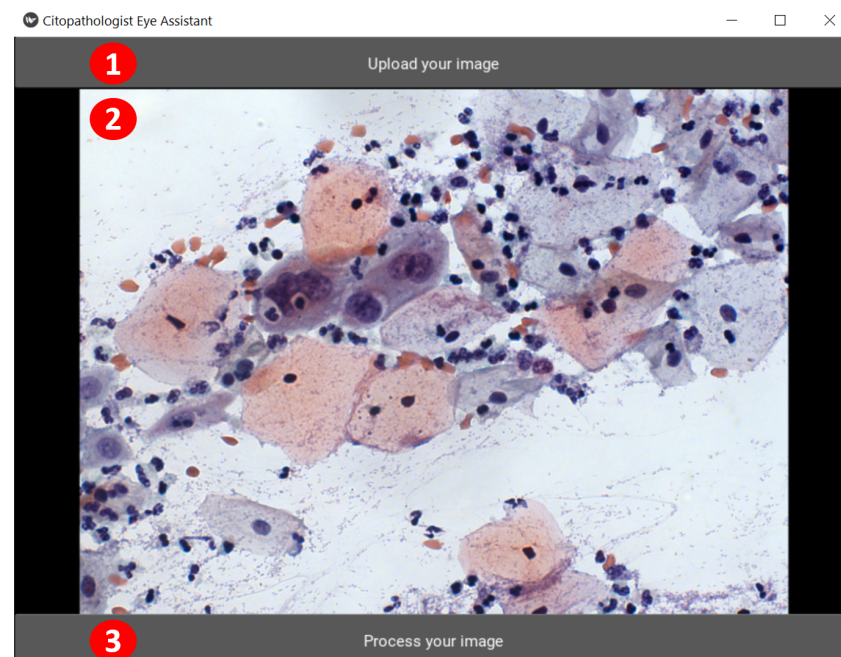


Figure 5. CEA home screen.

After the image processing step, the application presents the second screen, shown in Figure 6. CEA displays the resulting image at the position highlighted by the red circle 9. If the user wants to view the original image, he can click on the “no tags” button (see red circle 4). The “all” button (see red circle 5) displays the user-submitted image with all tags

found. The user can also select only lesioned cells marked (see red circle 6) or only negative cells marked (see red circle 7). In the position highlighted by red circle 8, the user will see the overall automatic classification of the image. Finally, in the “Process a new image” button (see red circle 10), the user can choose a new image for classification, returning to the first screen shown in Figure 5.

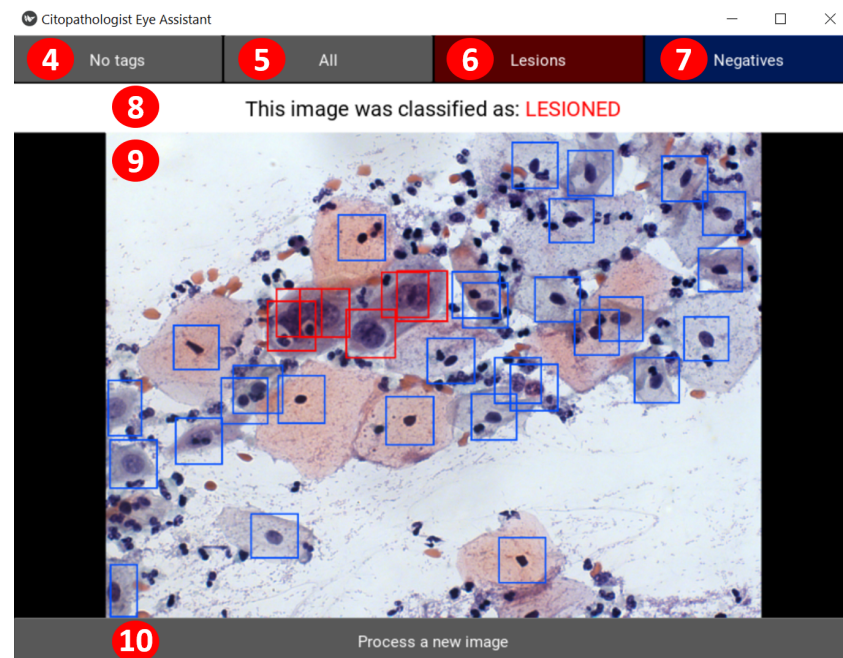


Figure 6. CEA evaluation screen.

5.1. CEA Validation Procedure

We designed a validation procedure to analyze the use of CEA and what specialists think about its applicability, possible improvements, or unnecessary functionalities.

The procedure was performed with three authors with 7, 12, and 24 years of experience in cytopathology. Moreover, for this evaluation, they were not aware of the construction of the tool or its characteristics.

Initially, the users (the authors performing the process) were asked to use the tool to classify four images. They have not received any prior instructions on how to use the tool. Then, the users were informed of the effect of “no tags”, “all”, “lesions”, and “negatives” buttons and notified that they could use them to assist their analysis. After that, the users classify another four images using CEA. In each image, the users answered if they agreed with the image’s overall classification and if the tags helped them in the analysis.

We gathered the images used for testing from the model’s test set. We separated the images from the model’s test set into four groups: only lesioned images, only negative images, images with few markings, and images with many markings and overlaps. We defined these groups because they represent the primary cases of application use in a more generic way. Thus, we randomly selected two images from each of these four groups for testing with users, totaling eight images.

After interacting with CEA, participants filled out a questionnaire in which they were asked if they thought the buttons helped them and if they thought the tool contributed to the image analysis. In addition, they evaluated the ease of use and learning, intuitiveness, and response time of the tool on a five-point Likert scale (1 as strongly disagree and 5 as strongly agree) based on their experience with the tool during the test. Finally, we conducted a semi-structured interview to gain more insights into the participants’ experiences.

5.2. Specialists' Feedback

During the test, users agreed with the image overall classification made by the CEA tool in 83% of the cases. They reported that some cases lack general context information of the slide (similar to other cells) to have confidence in the result. However, another user reported that these buttons could contribute to specific cases.

Specialists described that in some cases, the color, the number, and the overlapping of tags on the image might have negatively impacted the analysis since these characteristics may have influenced the user's ability to capture information. At this point, they reported that the filter buttons helped in the investigation. During the interview, a user suggested that the buttons for negatives filter and no tags are unnecessary, although it does not interfere. However, another user reported that buttons might not be needed in all analyses but can contribute to specific cases.

All users scored 5 (strongly agree) on the Likert scale concerning the statements analyzed in the test. That is, they agreed that the application was intuitive, easy to learn, and to use. During the interviews, users reinforced the application's ease of learning and used it as a positive point. Furthermore, they also rated the processing time of 8 seconds (on average) as satisfactory. However, during the interview, we realized that this evaluation was based on different reasons, varying from response time expectations and experience with other algorithms and tools.

During the interviews, some users highlighted that we could improve the flow of interaction with the system. They highlighted two distinct viewpoints: (1) remove non-essential steps; (2) change evaluation structure. For case (1), one user highlighted that viewing the image between loading it and submitting it allows them to perform a preliminary visual analysis of the image. However, this user understood that we only used it to confirm the chosen image. For case (2), the user highlighted that the process would become exhaustive and error-prone for multiple images due to repeated steps (select, evaluate, repeat). One suggestion in the interviews is that the tool could allow the loading and processing of multiple images.

Another observation reported by users concerns the number of classified classes. They confirmed that the classification into two classes is interesting for them, but they pointed out that it would be better to have more detail on the lesion class by classifying it into three or six classes according to the lesion type. Furthermore, they reported that it would be interesting if the application could handle images with different characteristics (such as zoom, resolution, and noise) since even similar equipment still generates slightly different images. However, we do not find a database with these characteristics. Thus, it is not possible yet to train a model capable of dealing with this variability.

Finally, users reported that the application could be beneficial in answering specific questions. One of these cases is when there is doubt and no other professional to assist them. Currently, the process adopted consists of taking a photo and sending it to other professionals.

6. Conclusions

This work presents the CEA (Cytopathologist Eye Assistant) tool that detects and classifies cells in multicellular images obtained in the Pap smear. Internally, the CEA processes utilize a detection algorithm and a classification algorithm to provide the results for the specialist. The application uses the YOLOv5s, a deep learning model for both tasks in this work. The goal is to support the cytopathologist's decision-making by offering a second opinion based on the whole image, not demanding a previous cell segmentation.

This work investigated different configurations for cervical cell detection and classification using YOLO architecture. The best results obtained were using the YOLOv5s with the original dataset. For this configuration, we obtained a precision of 0.736, recall of 0.777, and mAP@0.5 of 0.806. In addition, we evaluate the classification process as a separate task. For this analysis, we used the ensemble proposed by Diniz et al. [17] and obtained

a precision of 0.726 and a recall of 0.906 for classifying the bounding boxes detected by the YOLOv5s.

The analysis becomes even more interesting when we analyze the image as a whole. In this case, if the image has at least one lesioned cell, it is classified as lesioned. Otherwise, it is classified as negative. For this analysis, the YOLOv5s achieved the best results with 0.975 precision, 0.992 recall, 0.970 accuracy, and 0.983 F1-score for the CRIC Cervix database. The results outperform the literature method proposed by Xiang et al. [22] concerning the precision, recall, and F1 metrics; however, as they used liquid-based cytology, their images were easier to detect and classify than ours.

We can conclude from the specialists' assessment of CEA that it contributes to a general analysis of Pap smear images. They also reported that the tool is easy to learn and use, intuitive, and has an adequate response time. The main improvements suggested by users to be incorporated into the application were the following: (1) allowing the user to select multiple images for processing and (2) making a new version of the tool for classification into three or six classes.

Author Contributions: Conceptualization, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C., E.J.S.L., D.M.U., F.N.S.d.M. and M.J.F.S.; methodology, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C. and M.J.F.S.; software, D.N.D. and B.N.S.K.; validation, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C., R.R.e.R.O. and M.J.F.S.; formal analysis, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C., R.R.e.R.O., E.J.S.L. and M.J.F.S.; investigation, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C., D.M.U., F.N.S.d.M. and M.J.F.S.; resources, A.G.C.B., C.M.C. and M.J.F.S.; data curation, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C., R.R.e.R.O. and M.J.F.S.; writing—original draft preparation, D.N.D., B.N.S.K. and M.T.R.; writing—review and editing, D.N.D., B.N.S.K., M.T.R., A.G.C.B., C.M.C., E.J.S.L., D.M.U., F.N.S.d.M. and M.J.F.S.; visualization, D.N.D., B.N.S.K. and M.T.R.; supervision, A.G.C.B., C.M.C., E.J.S.L. and M.J.F.S.; project administration, A.G.C.B., C.M.C. and M.J.F.S.; funding acquisition, A.G.C.B., C.M.C., E.J.S.L. and M.J.F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded 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—PROPPI/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].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this work is available at CRIC Searchable Image Database (<https://database.cric.com.br>) (accessed on 1 November 2022).

Acknowledgments: The authors would like to thank the Universidade Federal de Ouro Preto (UFOP), the Center for Recognition and Inspection of Cells (CRIC), the Extended Reality for Good Laboratory (XR4Good), the Group of Optimization and Algorithms (GOAL), the Intelligent Systems Computing Laboratory (CSILab), the *Laboratório Multiusuários de Microscopia Avançada e Microanálise do Núcleo de Pesquisas em Ciências Biológicas (NUPEB)* and the Center for Advanced Mathematics for Energy Research Applications (CAMERA) for also supporting this research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ASC-H	Atypical Squamous Cells, cannot exclude HSIL
ASC-US	Atypical Squamous Cells of Undetermined Significance
CAD	Computer-aided diagnosis

CAMERA	Center for Advanced Mathematics for Energy Research Applications
CEA	Cytopathologist Eye Assistant
CNN	Convolutional Neural Network
CRIC	Center for Recognition and Inspection of Cells
CSILab	Intelligent Systems Computing Laboratory
FN	False Negative
FP	False Positive
GOAL	Group of Optimization and Algorithms
HSIL	High-grade Squamous Intraepithelial Lesion
LSIL	Low-grade Squamous Intraepithelial Lesion
mAP	Mean Average Precision
NILM	Negative for Intraepithelial Lesion or Malignancy
NUPEB	Núcleo de Pesquisas em Ciências Biológicas
SCC	Squamous Carcinoma
TIS	Thin Prep Imaging System
TP	True Positives
UFOP	Universidade Federal de Ouro Preto
UFOP	Universidade Federal do Ceará
XR4Good	Extended Reality for Good Laboratory
YOLO	You Only Look Once

References

- Chandrasekhar, V.; Krishnamurti, C. George Papanicolaou (1883–1962): Discoverer of the Pap Smear. *J. Obstet. Gynecol. India* **2018**, *68*, 232–235. [\[CrossRef\]](#)
- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J. Clin.* **2021**, *71*, 209–249. [\[CrossRef\]](#)
- Gay, J.; Donaldson, L.; Goellner, J. False-negative results in cervical cytologic studies. *Acta Cytol.* **1985**, *29*, 1043–1046.
- Bosch, M.; Rietveld-Scheffers, P.; Boon, M. Characteristics of false-negative smears tested in the normal screening situation. *Acta Cytol.* **1992**, *36*, 711–716.
- Naryshkin, S. The false-negative fraction for Papanicolaou smears. *Arch. Pathol. Lab. Med.* **1997**, *121*, 270.
- Franco, R.; Amaral, R.G.; Montemor, E.B.L.; Montis, D.M.; Morais, S.S.; Zeferino, L.C. Factors associated with false-negative cervical cytopathological results. *Rev. Bras. Ginecol. Obs.* **2006**, *28*, 479–485. [\[CrossRef\]](#)
- Amaral, R.G.; Manrique, E.J.C.; Guimarães, J.V.; Sousa, P.J.d.; Mignoli, J.R.Q.; Xavier, A.d.F.; Oliveira, A. Influência da adequabilidade da amostra sobre a detecção das lesões precursoras do câncer cervical. *Rev. Bras. Ginecol. Obs.* **2008**, *30*, 556–560. [\[CrossRef\]](#)
- Silva, G.P.F.; Cristovam, P.C.; Vidotti, D.B. O impacto da fase pré-analítica na qualidade dos esfregaços cervicovaginais. *Rev. Bras. Anal. Clin.* **2017**, *49*, 135–140.
- Koonmee, S.; Bychkov, A.; Shuangshoti, S.; Bhumichitra, K.; Himakhun, W.; Karalak, A.; Rangdaeng, S. False-negative rate of papanicolaou testing: A national survey from the Thai society of cytology. *Acta Cytol.* **2017**, *61*, 434–440. [\[CrossRef\]](#)
- Siddegowda, R.B.; DivyaRani, M.; Natarajan, M.; Biligi, D.S. Inter-Observer Variation in Reporting of Pap Smears. *Natl. J. Lab. Med.* **2016**, *5*, 22–25.
- Wilbur, D.; Nayar, R. Bethesda 2014: Improving on a paradigm shift. *Cytopathology* **2015**, *26*, 339–342. [\[CrossRef\]](#)
- Baena, A.; Guevara, E.; Almonte, M.; Arias-Stella, J.; Sasieni, P.; Sanchez, G. Factors related to inter-observer reproducibility of conventional Pap smear cytology: A multilevel analysis of smear and laboratory characteristics. *Cytopathology* **2017**, *28*, 192–202. [\[CrossRef\]](#)
- Branca, M.; Longatto-Filho, A. Recommendations on quality control and quality assurance in cervical cytology. *Acta Cytol.* **2015**, *59*, 361–369. [\[CrossRef\]](#)
- Rezende, M.T.; Bianchi, A.G.; Carneiro, C.M. Cervical cancer: Automation of pap test screening. *Diagn. Cytopathol.* **2021**, *49*, 559–574. [\[CrossRef\]](#)
- Rehman, A.u.; Ali, N.; Taj, I.; Sajid, M.; Karimov, K.S. An Automatic Mass Screening System for Cervical Cancer Detection Based on Convolutional Neural Network. *Math. Probl. Eng.* **2020**, *2020*, 4864835. [\[CrossRef\]](#)
- Ghoneim, A.; Muhammad, G.; Hossain, M.S. Cervical cancer classification using convolutional neural networks and extreme learning machines. *Future Gener. Comput. Syst.* **2020**, *102*, 643–649. [\[CrossRef\]](#)
- Diniz, D.N.; Rezende, M.T.; Bianchi, A.G.C.; Carneiro, C.M.; Luz, E.J.S.; Moreira, G.J.P.; Ushizima, D.M.; de Medeiros, F.N.S.; Souza, M.J.F. A Deep Learning Ensemble Method to Assist Cytopathologists in Pap Test Image Classification. *J. Imaging* **2021**, *7*, 111. [\[CrossRef\]](#)
- Diniz, D.N.; Rezende, M.T.; Bianchi, A.G.C.; Carneiro, C.M.; Ushizima, D.M.; de Medeiros, F.N.S.; Souza, M.J.F. A Hierarchical Feature-Based Methodology to Perform Cervical Cancer Classification. *Appl. Sci.* **2021**, *11*, 4091. [\[CrossRef\]](#)

19. Di Ruberto, C.; Loddo, A.; Putzu, L. Histological image analysis by invariant descriptors. In *Image Analysis and Processing, Proceedings of the International Conference on Image Analysis and Processing (ICIAP 2017), Catania, Italy, 11–15 September 2017*; Battiato, S., Gallo, G., Schettini, R., Stanco, F., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2017; Volume 10484, pp. 345–356.
20. Bora, K.; Chowdhury, M.; Mahanta, L.B.; Kundu, M.K.; Das, A.K. Automated classification of Pap smear images to detect cervical dysplasia. *Comput. Methods Programs Biomed.* **2017**, *138*, 31–47. [[CrossRef](#)]
21. Matias, A.V.; ao Gustavo Atkinson Amorim, J.; Macarini, L.A.B.; Cerentini, A.; Onofre, A.S.C.; Onofre, F.B.D.M.; Daltoé, F.P.; Stemmer, M.R.; von Wangenheim, A. What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review. *Comput. Med. Imaging Graph.* **2021**, *91*, 101934. [[CrossRef](#)]
22. Xiang, Y.; Sun, W.; Pan, C.; Yan, M.; Yin, Z.; Liang, Y. A novel automation-assisted cervical cancer reading method based on convolutional neural network. *Biocybern. Biomed. Eng.* **2020**, *40*, 611–623. [[CrossRef](#)]
23. Bengtsson, E.; Malm, P. Screening for cervical cancer using automated analysis of PAP-smears. *Comput. Math. Methods Med.* **2014**, *2014*, 842037. [[CrossRef](#)]
24. Thrall, M.J. Automated screening of Papanicolaou tests: A review of the literature. *Diagn. Cytopathol.* **2019**, *47*, 20–27. [[CrossRef](#)]
25. Hologic, Inc. Genius™ Digital Diagnostic System. 2021. Available online: https://www.hologic.com/file/51136/download?token=S7A_JhgH (accessed on 31 October 2022).
26. Crowell, E.F.; Bazin, C.; Saunier, F.; Brixtel, R.; Caillot, Y.; Lesner, B.; Toutain, M.; Ferreri, C.; Garcia, I.; Mathieu, M.C.; et al. CytoProcessorTM: A New Cervical Cancer Screening System for Remote Diagnosis. *Acta Cytol.* **2019**, *63*, 215–223. [[CrossRef](#)]
27. Crowell, E.F.; Bazin, C.; Thurotte, V.; Elie, H.; Jitaru, L.; Olivier, G.; Caillot, Y.; Brixtel, R.; Lesner, B.; Toutain, M.; et al. Adaptation of CytoProcessor for cervical cancer screening of challenging slides. *Diagn. Cytopathol.* **2019**, *47*, 890–897. [[CrossRef](#)]
28. Wang, C.Y.; Liao, H.Y.M.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. CSPNet: A new backbone that can enhance learning capability of CNN. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA, 14–19 June 2020; pp. 1571–1580.
29. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.
30. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path Aggregation Network for Instance Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.
31. Wang, C.Y.; Yeh, I.H.; Liao, H.Y.M. You Only Learn One Representation: Unified Network for Multiple Tasks, 2021. *arXiv* **2021**, arXiv:2105.04206.
32. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
33. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
34. Jocher, G. YOLOv5. 2020. Available online: <https://github.com/ultralytics/yolov5> (accessed on 29 July 2021).
35. Kathuria, A. How to Train YOLO v5 on a Custom Dataset. 2021. Available online: <https://blog.paperspace.com/train-yolov5-custom-data/> (accessed on 30 July 2021).
36. Rezende, M.T.; Silva, R.; Bernardo, F.D.O.; Tobias, A.H.; Oliveira, P.H.; Machado, T.M.; Costa, C.S.; Medeiros, F.N.; Ushizima, D.M.; Carneiro, C.M.; et al. Cric searchable image database as a public platform for conventional pap smear cytology data. *Sci. Data* **2021**, *8*, 1–8. [[CrossRef](#)]
37. Casado-García, Á.; Domínguez, C.; García-Domínguez, M.; Heras, J.; Inés, A.; Mata, E.; Pascual, V. CLoDSA: A tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks. *BMC Bioinform.* **2019**, *20*, 323. [[CrossRef](#)]
38. Dhurandhar, A.; Dobra, A. Insights into Cross-Validation. 2008. Available online: <https://s3.us.cloud-object-storage.appdomain.cloud/res-files/2201-CrossVnDMKD.pdf> (accessed on 30 July 2021).
39. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How Transferable Are Features in Deep Neural Networks? In *Proceedings of the 27th International Conference on Neural Information Processing Systems—Volume 2*, Montreal, QC, Canada, 8–13 December 2014; pp. 3320–3328.
40. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, 22–24 June 2009; pp. 248–255.
41. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In *Proceedings of the European Conference on Computer Vision*, Zurich, Switzerland, 6–12 September 2014; Bertino, E., Gao, W., Steffen, B., Yung, M., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8693, pp. 740–755.