Nucleus Detection in Cervical Samples Stained With AgNOR

João Gustavo Atkinson Amorim
joao.atkinson@posgrad.ufsc.br
Federal University of Santa Catarina
Florianópolis, Santa Catarina, Brazil

Vinícius Moreno Sanches
Federal University of Santa Catarina
Florianópolis, Santa Catarina, Brazil
vinicius.moreno.sanches@grad.ufsc.br

Tainee Bottamedi
Federal University of Santa Catarina
Florianópolis, Santa Catarina, Brazil
tainee.bottamedi@grad.ufsc.br

André Victória Matias
Federal University of Santa Catarin
Florianópolis, Santa Catarina, Brazi
andre.v.matias@posgrad.ufsc.br

Marco Antônio Martins Cavaco
Federal University of Santa Catarina
Florianópolis, Santa Catarina, Brazil
m.cavaco@ufsc.br

Alexandre Sherlley Onofre
Federal University of Santa Catarin
Florianópolis, Santa Catarina, Brazi
alexandre.onofre@ufsc.br

Fabiana Botelho Onofre
Federal University of Santa Catarin
Florianópolis, Santa Catarina, Brazi
fabiana.onofre@ufsc.br

Aldo Von Wangenheim
Federal University of Santa Catarin
Florianópolis, Santa Catarina, Brazi
aldo.vw@ufsc.br

ABSTRACT

Cervical cancer is a public health problem, where the treatment has a better chance of success if detected early. This paper explores one way of to analyze argyrophilic nucleolus organizer regions (AgNOR) stained slide using deep learning approaches of object detection for detecting the different categories of nucleus. Our results show that a balanced dataset between the explored categories was essential, also that a ResNet-50 as backbone of Fast RCNN shows an AP of 61.8% and 42.5% to detect nucleus and out of focus nucleus.

KEYWORDS

Cytology; AgNOR; Deep Learning;

1 INTRODUCTION

Cervical cancer, predominantly caused by persistent infection by some types of Human Papillomavirus (HPV) called oncogenic. It is characterized by the disordered replication of the organ lining epithelium, which can invade other structures until it reaches circulation and affect other parts of the body [1]. HPV is the most common viral infection of the reproductive tract. Two types (16 and 18) cause 70% of cervical cancers and precancerous lesions [2]. This disease prevention can be done in several ways, but it is still one of the main causes of death from cancer worldwide. Today, it is the seventh most incident type of cancer, with approximately 570 thousand new cases per year, being the fourth most common cancer among women [3]. In Brazil, the National Cancer Institute (2020) [4] estimates that more than 16 thousand new cases will occur for each year of the 2020-2022, occupying the third position of incidence in women across the country. Furthermore, the diagnosis made at an advanced stage is one of the most critical reasons why cervical cancer maintains high mortality rates all over the world [5].

Usually, cervical cancer can be tracked early by detecting precancerous changes, and it can present a favorable prognosis, with great chances of cure. Currently, the most used screening tests are cervical cytology (or Pap test), visual inspection with acetic acid and the HPV test for high-risk types [6]. Depending on the smear test results, the patient is referred to colonoscopy and a biopsy may be necessary in more extreme cases for an accurate evaluation of the lesion. These conventional techniques, although quite reliable, have low rates of reproducibility and are not so cheap and effective. Moreover, they present high rates of false negatives. Because of that, researchers see the need to adopt new cytological techniques for predictive, reliable, and reproducible analyzes [7].

Thus, an alternative way of investigating cervical samples is the detection of nucleolus organizer regions (NORs). These are DNA loops containing genes responsible for the transcription of ribosomal RNA located in the cell nucleolus. They contain a set of argyrophilic proteins, selectively stained by silver nitrate, which can be identified as black dots located throughout the nucleoli area and called AgNORs [8]. The numbers of AgNOR in interphase are closely related to cell proliferative activity, since it is the region of the nucleus where ribosomal RNA's are synthesized, processed, and assembled with ribosomal proteins. Therefore, research on these biological markers has expressed prognostic importance since it show that malignant cells often have higher numbers of AgNOR compared to corresponding benign or normal cells [9]. The advantages of the AgNOR technique over other methods of evaluating cell proliferation are simplicity and economy, since it can be applied to a material routinely processed, the reagents can be easily found and have a reduced cost [10]. The biggest disadvantage, although it presents results with a low probability of error, is the manual evaluation. There is some risk of variations caused by the observation conditions and the great propensity to human errors during the process. In addition, the technique demands a lot of time, which also occurs with other types of dyes [11].

To minimize human errors during this diagnostic process, Machine Learning methods are an adequate alternative to provide more reliable results and in greater quantity. Within this perspective, neural networks have been used for more than 25 years to perform quantitative analyzes on samples of cytological material [12]. Recently, however, due to the great growth around computer vision, Deep Learning algorithms have been achieving spectacular

results in the mission of detecting the most varied types of objects, bringing the possibility of extracting more useful information from data [13]. For this reason, they can be used to assist in the identification and classification of nucleolus organizing regions by cytological analysis [14].

This paper presents the results of applying object detection methods to silver-stained cytology images (AgNOR technique), using deep learning methods for computer vision. The main objective was to detect nuclei and separate them into different categories.

2 RELATED WORKS

A systematic literature review made in middle 2020 with the objective of finding what are the current computer-assisted or artificial intelligence-based approaches in computer vision for the support of quantitative cytology and diagnosis of cancer in cytological exams [14]. The studies analyzed was from the beginning of 2016 to middle of 2020. Based on the analysis of the papers, deep learning approaches (mostly CNNs) were the most common in these works. However, classical approaches such as Support Vector Machine, Random Forest, Artificial Neural Network, super-pixel segmentation, and others, are still employed in some works.

Due to the various problems encountered with automatic screening caused by improper staining of cells and overlapping cell images, a methodology proposed in [15] to deal with the problem in nucleus segmentation and to suggest classification methods for automatic screening of cervical cancer. For the segmentation of the nucleus, it was applied k-means clustering [16] at the pre-processed images. For the classification, it was used random forest classifier, after comparing the results obtained in other classification techniques. Similar works, [17] and [18] were performed, showing the wide range of projects that seek to apply computational methods to minimize human error in the process of cancer detection.

In [19] the authors aimed to investigate the diagnostic value of AgNOR counts in cervical smears and discriminate the degrees of squamous intraepithelial lesion (SIL). Results demonstrated a progressive increase in AgNOR counts when the severity of the lesions increased, mainly in the differentiation between low and high-grade SIL. In addition, on [20] and [21] they analyzed the AgNOR count as a marker of malignancy progression. As the count shows the intensity of cell proliferation, this method can be used in differentiating between low and high degree of cervical dysplasia, as well as between lesions with high and low potential for malignant transformation. The technique presented prognostic value in the early identification of high-risk cases, as can be seen in [22].

Lastly, in [12], the authors used a pipeline for AgNOR quantification, focused on cytology exams using deep learning algorithms aiming at the early detection of cervical neoplasm. They used a U-Net with ResNet18 as the backbone to perform the semantic segmentation. Moreover, with their results on the segmentation process, the quantity, and area of the NORs were measured. This gives us some conclusions about how to early detect cancer using AgNORs. In this work, the best result was a 0.87 of restrict mIoU and 0.99 of restrict DSC. These results validate the use of semantic segmentation to measure AgNORs.

3 MATERIAL AND METHODS

3.1 Dataset

The dataset has been built with the Gynecology and Colonoscopy Outpatient Clinic of the University Hospital Professor Polydoro Ernani de São Thiago of Federal University of Santa Catarina (HU UFSC EBSERH). It was built by images from examinations performed on women that had presented cytological alterations in the gynecological cytology of previous exams. Presented in [12] a first version of the CCAgT dataset, with a total of 2540 images from slides of 3 patients. The present work use an intermediate version (before second version of CCAgT) containing new categories for nuclei. This intermediary version has a total of 3905 images from slides of 5 patients, with a total of 30.912 annotations, and the distribution of the categories is in the Table 1.

Table 1: Division of annotations by categories

Name	Number of annotations	%
Nucleus	7509	24
Cluster	15074	49
Satellites	3913	13
Nucleus out of focus	3177	10
Overlapped nuclei	628	2
Non-viable nucleus	611	2

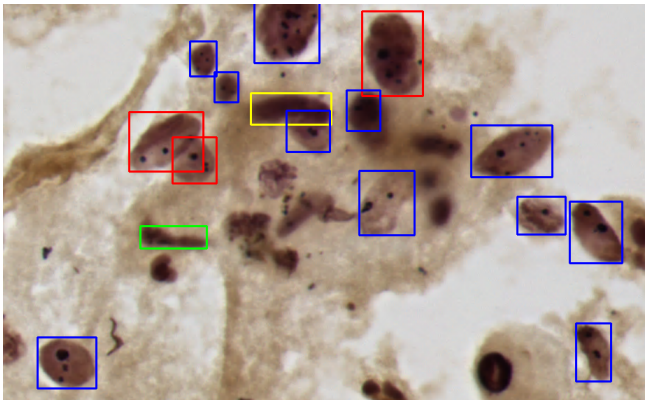
Samples of the dataset can be seen in Fig. 1, being Fig. 1a a sample of a tile/image with the nuclei categories highlighted by the boundary boxes. The Fig. 1b shows samples of each nucleus category. The colors of the boundary boxes represents each category, these being:

- Nucleus: in blue, that reefers to an object that contains NORs and can be analyzed/diagnosed;
- Overlapped nuclei: in red are the nuclei that was overlapped, which cannot be analyzed because it is not possible to identify to which nuclei the NOR belongs;
- Non-viable nucleus: in green are the samples deformed in a way that does not allow the analysis of it;
- Nucleus out of focus: in yellow the samples that reefers to a nucleus out of focus, in a way that does not allow its internal NORs to be identified.

3.2 Models

The Fast Region-based Convolution Network [23] is a network that, by an image input, and a set of objects, processes the whole image with max polling layers and several convolutions to achieve a convolution feature map. Subsequently, this network takes for each object, is extracted a fixed-length feature vector from the feature map by a region of interest (RoI) pooling layer. These RoI Feature Vectors results in two outputs: a soft max probability, and a bounding-box regression.

With the purpose of detection of the four categories of nucleus, has been tested different models of Fast R-CNN. The backbone tested models were ResNet-50 (R50), R50-C4, ResNet-101 (R101) and ResNeXt-101-32x8d (X101), was listed in table 2 all the models.



(a) Sample of an image from CCAgT dataset

(b) Samples of each nucleus category from CCAgT dataset

Figure 1: Sample of CCAgT dataset

Model R50-C4 being to an FPN (with ResNet-50) that extracted features from the final layer of 4th stage convolutions. These models trained with 1x or 3x of the learning rate schedule, where 1x refers to approximately 12 COCO epochs. The entire training and testing process conducted in the Python 3.7.11 with PyTorch [24] at version 1.7.1, and the Detectron2 [25] at version 0.5.

Table 2: The Fast R-CNN tested models

Name	Backbone	learning rate schedule
R50-1x	ResNet-50	1x
R50-1x-C4	ResNet-50-C4	1x
R50-3x-C4	ResNet-50-C4	3x
R101-3x	ResNet-101	3x
X101-3x	ResNeXt-101-32x8d	3x

In general, the models have similar characteristics, being capable of to be trained using the same parameters. However, each one has different depths and quantities of layers along the network. Therefore, the experiments in the models referred has the purpose of evaluate different architectures and determine which one has the best performance when finding nucleus and stained images with AgNOR.

Besides being state-of-the-art models, they are already widespread in the literature, and no need for further explanations on the details of each model because this is an application paper. All models implemented with Detectron2 and PyTorch at background, and the detailing of each tested architecture explained in [25].

3.3 Metrics

As demonstrated and discussed in [14], one of the most important factors about the training of the models is to avoid biased results. Considering it was split the dataset into three different groups (train, validation, and test set), where each fold:

- **Training set** is the subject used for training the model;
- **Validation set** also used during the training of the model, but its hole is to test the network, besides be used for parameter tuning;
- **Test set** is the subject used after the training, to measure how is the performance of the network. This dataset is the one used to calculate the metrics described next.

As the goal of this paper is object detection, will be used the Average Precision (AP) to measure the performance of the networks. Furthermore, according to [14], the average precision is a convenient metric for object detection. Calculated by the area under the Precision-Recall curve, it is, compute for each image of the test set the precision and recall.

As discussed in [26] and in [27], the precision (P) (Equation 1) is calculated by the ratio between the True Positives (TP) and the sum of False Positives (FP) and True Positives. While recall (R) (Equation 2) is the ratio between the True Positives and the sum of False Negatives (FN) and the True Positives.

$$P = \sum_{j=1}^{k} \frac{n_{jj}}{n_{ij} + n_{jj}}, \quad i \neq j. \tag{1}$$

$$R = \sum_{j=1}^{k} \frac{n_{jj}}{n_{ji} + n_{jj}}, \quad i \neq j. \tag{2}$$

Where n_{xy} represents the number of pixels classified as x, and labeled as y.

3.4 Experiment setup

The **first experiment** (E1) was made using all annotations of nuclei available. Adopting a learning rate (LR) of 5×10^{-4}, with weight decay (WD) of 5×10^{-4}, and was performed the training during 1200 iterations (iter).

In the **second experiment** (E2), it was balanced the number of annotations for each class in the dataset, looking to analyze the variation of results by the balance of the nuclei categories of the dataset. The balance of the dataset randomly done, that is, it was deleted random annotations of each class. In addition, the images without any categories also were deleted. The balanced annotation results in 1534, 425, and 428 annotations of nuclei, out of focus, overlapped, and non-viable respectively, using the same hyperparameters as E1.

For the **third experiment** (E3), was made a hard balance of the nuclei categories, where reduced the number of annotations to 706 Nuclei, 632 Out of Focus, 419 Overlapped, and 404 Non-viable nuclei. Furthermore, increased the iterations number to 2000 without overfitting the model.

The **fourth experiment** (E4) was performed using the same setup as E5, but at E4 was applied a pipeline of data augmentation

Table 3: Experiment setups

Experiment	Nº of annotations				LR	WD	iter	previous setup
	Nucleus	Out of focus	Overlapped	non-viable				
E1	5193	2247	435	404	5×10^{-4}	5×10^{-4}	1000	-
E2	1534	678	425	428	5×10^{-4}	5×10^{-4}	1000	-
E3	706	632	419	404	5×10^{-4}	5×10^{-4}	2000	-
E4	706	632	419	404	5×10^{-4}	3×10^{-4}	2000	-
E5	706	632	419	404	5×10^{-4}	3×10^{-4}	2000	pre-train the model with 3624 nuclei samples
E6	2932	2225	-	-	5×10^{-4}	5×10^{-4}	2500	pre-train the model with 1936 nuclei samples

at the training step. The data augmentation consists in apply random brightness, contrast, saturation, rotation, flip, and crop with a probability of 50%.

At the **fifth experiment** (E5), was made a transfer learning technique with the same dataset. So, first was trained the model for 2000 iterations just for nuclei detection, with 3624 annotations deleted previously. After this, were transferred the parameters and the model has trained again with just the balanced dataset created at E3.

Lastly, the **sixth experiment** (E6) was trained the model just to detect nucleus and out-of-focus nucleus, that was the categories of nuclei that have a higher quantity of annotations. At this experiment, first was adopted the transfer learning with the samples of nuclei that isn't at the training set. With that, first, train the model for just nuclei detection using 2261 nuclei annotations. So, train the model using 2932 and 2225 annotations of nuclei and nuclei out of focus, respectively. Performing the train with a learning rate of 5×10^{-4}, with a weight decay of 5×10^{-4}, and 2500 iterations.

In both experiments, E5 and E6 were pre-trained with a learning rate of 5×10^{-4}, with a weight decay of 5×10^{-4}, and 2000 iterations. The summary of experiments setups was in table 3.

4 RESULTS AND DISCUSSION

The first experiment performed followed the details of **E1**, for the five models chosen. The table 4 show the results for each model and category of nucleus. From these results has been determined the model that best perform for all categories to follow with the rest of experiments.

The best nucleus and nucleus out of focus detection at E1 has with the model R101-3x with 61.3% and 28.4% AP respectively. For overlapped nuclei and non-viable nucleus, all models underperform, with the R50-1x being the unique successfully to detect some overlapped nucleus. The models R50-1x-C4, R50-3x-C4 and C101-3x does not present relevant results, without AP higher than 2% for the out of focus, overlapped and non-viable nucleus. Moreover, present lower values of AP for nucleus detection.

The R50-1x performs similar to the R101-3x, but with better results for overlapped nucleus. Because it is a smaller model, had a similar result to R101-3x, and was able to perform moderately for the overlapping nuclei, we set it as the ideal model to be tested in the remaining experiments. For the E1, the R50-1x had an AP of 58.2% and 27.5% for nucleus and nucleus out of focus respectively, also 2.5% and 2.2% for overlapped and non-viable nucleus. Although they are not good results for these two categories, understood that the responsibility of under-perform is the quantity of samples.

Table 4: Results for each models at E1

Model	Avg. Precision (%)			
	Nucleus	Out of focus	Overlapped	non-viable
R50-1x	58.2	27.5	2.5	2.2
R50-1x-C4	10.4	2.0	0	0.1
R50-3x-C4	10.0	1.2	0	0.3
R101-3x	61.3	28.4	0.1	2.8
X101-3x	7.6	0	0	0

Looking for a better set of training for better results, the other experiments have been conduct and the summarized results can be seen at table 5. As expected, when conduct a balancing of the sample as in E2 improve the performance for the out of focus, overlapped and non-viable nucleus without under-performing for the nucleus category.

To improve the results in general, the E3, E4 and E5 use the same samples with a hard balancing seeking the best possible balance between the categories. The augmentations applied in E4 shows the model under-perform in comparison with the results of E3. The E3 results shows an equilibrium between the AP for each category, but the higher results have been achieved by the E5 showing that the transfer learning can boost the results for nuclei detection.

Comparing the results between the E3 and E5, shows an improvement of 11.8% and 3.2% at AP for out of focus and non-viable nucleus at E5. For nucleus and overlapped nuclei detection, the E3 shows higher in 3.6% and 3.9% of AP compared with E5. By the increases achieved by the E5 is higher than E3, the best model for multicategory detection of AgNOR nucleus was the model performed at E5. Although lowering the quantity of samples for the E3, E4 and E5 when compared with E1 and E2, shows that the higher balance dataset for multicategory detection achieved better performance.

Furthermore, the results show that a balanced dataset is most important for this work than the pipeline of data augmentation

applied for the multicategory nucleus detection. Looking to extract the better cases of this version of dataset, was performed E6 to detection of nucleus and out of focus nucleus. At E6, that consists on a hard balance between in just the two categories, increasing the number of samples when compared with E3. This approach shows the highers values of AP compared with all experiments with an 61.8% and 42.5% AP for nucleus and out of focus nucleus respectively. The results of E6, is the result of a larger number of samples with the best experiment found (E5).

Table 5: Results for experiments with Faster RCNN with ResNet-50 (R50-1x)

Experiment	Avg. Precision (%)			
	Nucleus	Out of focus	Overlapped	non-viable
E2	56	30	9	19
E3	34.6	26.2	16.2	15.8
E4	18.6	9.8	7.8	9.8
E5	31	38	12.3	19
E6	61.8	42.5	-	-

5 CONCLUSION

In this work, we evaluated models of deep learning for object detection at an intermediary version of CCAgT dataset. Tests different approaches to detect multicategories of nucleus from cervical samples stained with AgNOR technique. The better model as backbone for a Fast RCNN is the ResNet-50 in this work, but also the ResNet-101 can be explored in the future.

Among the experiments performed the use of ResNet-50 as backbone, with a balanced dataset and using the rest of nucleus samples for pre-train the model shows the best performance for multicategorical detection. This experiment got the AP of 31%, 38%, 12.3% and 19% for nucleus, nucleus out of focus, overlapped nuclei and non-viable nucleus respectively. Because these are experiments at intermediate dataset, this work shows exciting results for the detection of nuclei stained with Agnor, although the model does not obtain high AP values.

When testing this pipeline of experiment, with just the nucleus and nucleus out of focus, that have higher quantities of samples, the AP increases significantly. This points that this pipeline or this approach can achieve better results when the model receive more samples. For the two categories, the Fast RCNN with ResNet-50 have an AP of 61.8% and 42.5% for nucleus and out of focus respectively.

In future works, we want to test different methods (as semantic segmentation and instance segmentation) to build a complete pipeline, also test a cascade model with object detection and semantic segmentation to automated slide analysis to help in the early diagnosis and prognostic of cervical cancer cases. Furthermore, we want to enlarge the size (in image and patients quantities) of the dataset, to help improve the model's assertiveness.

REFERENCES

[1] Instituto Nacional de Câncer. Câncer do colo do útero, 2018. URL https://www.inca.gov.br/tipos-de-cancer/cancer-do-colo-do-utero. (accessed: 28-02-2021).

[2] Janete Tamani Tomiyoshi Nakagawa, Janine Schirmer, and Márcia Barbieri. Vírus HPV e câncer de colo de útero. *Revista Brasileira de Enfermagem*, 63(2):307–311, April 2010. doi: 10.1590/s0034-71672010000200021. URL https://doi.org/10.1590/s0034-71672010000200021.

[3] J Ferlay, M Ervik, F Lam, M Colombet, L Mery, M Piñeros, A Znaor, I Soerjomataram, and F Bray. Global cancer observatory: Cancer today, 2020. URL https://gco.iarc.fr/today. (accessed: 28-02-2021).

[4] Instituto Nacional de Câncer. Conceito e magnitude, 2018. URL https://www.inca.gov.br/controle-do-cancer-do-colo-do-utero/conceito-e-magnitude. (accessed: 28-02-2021).

[5] Luiz Claudio Santos Thuler and Gulnar Azevedo Mendonça. Estadiamento inicial dos casos de câncer de mama e colo do útero em mulheres brasileiras. *Revista Brasileira de Ginecologia e Obstetrícia*, 27(11):656–660, November 2005. doi: 10.1590/s0100-72032005001100004. URL https://doi.org/10.1590/s0100-72032005001100004.

[6] Inc. American Cancer Society. Can cervical cancer be prevented?, 2020. URL https://www.cancer.org/cancer/cervical-cancer/causes-risks-prevention/prevention.html. (accessed: 28-02-2021).

[7] Carla Filippin, Larissa Duarte Chistofoletti, Maria Cecília Menks Ribeiro, and CI VITURI. Determinação do número de regiões organizadoras de nucléolo (agnor) em lesões do epitélio cervical uterino. *Revista Brasileira de Análises Clínicas, Rio de Janeiro*, 38(3):133–139, 2006.

[8] Flávio de Oliveira Lima, Regina Maria Catarino, Celina Tizuko Fujiyama Oshima, and Marco Túllio de Assis Figueiredo. Regiões organizadoras nucleolares argirofilicas no sarcoma sinovial. *Jornal Brasileiro de Patologia e Medicina Laboratorial*, 41(5):347–352, 2005.

[9] D Trerè. Agnor quantification in tumour pathology: what is actually evaluated? *Journal of clinical pathology*, 46(2):189, 1993.

[10] M Derenzini. The AgNORs. *Micron*, 31(2):117–120, April 2000. doi: 10.1016/s0968-4328(99)00067-0.

[11] Patrícia Campos Fontes. Avaliação quantitativa das proteínas agnors na citologia esfoliativa da mucosa bucal de pacientes fumantes e não-fumantes. *BBO - Odontologia*, 2005.

[12] João Gustavo Atkinson Amorim, Luiz Antonio Buschetto Macarini, André Victória Matias, Allan Cerentini, Fabiana Botelho De Miranda Onofre, Alexandre Sherlley Casimiro Onofre, and Aldo von Wangenheim. A novel approach on segmentation of agnor-stained cytology images using deep learning. pages 552–557, 2020. doi: 10.1109/CBMS49503.2020.00110. URL https://doi.org/10.1109/CBMS49503.2020.00110.

[13] Melih Kandemir and Fred A. Hamprecht. Computer-aided diagnosis from weak supervision: A benchmarking study. *Computerized Medical Imaging and Graphics*, 42:44–50, June 2015. doi: 10.1016/j.compmedimag.2014.11.010. URL https://doi.org/10.1016/j.compmedimag.2014.11.010.

[14] André Victória Matias, João Gustavo Atkinson Amorim, Luiz Antonio Buschetto Macarini, Allan Cerentini, Alexandre Sherlley Casimiro Onofre, Fabiana Botelho De Miranda Onofre, Felipe Perozzo Daltoé, Marcelo Ricardo Stemmer, and Aldo von Wangenheim. What is the state of the art of computer vision-assisted cytology? a systematic literature review. Technical report, July 2021. URL https://doi.org/10.1016/j.compmedimag.2021.101934.

[15] Hmrishav Bandyopadhyay and Mita Nasipuri. Segmentation of pap smear images for cervical cancer detection. In *2020 IEEE Calcutta Conference (CALCON)*, pages 30–33, 2020. doi: 10.1109/CALCON49167.2020.9106484.

[16] G.B. Coleman and H.C. Andrews. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979. doi: 10.1109/PROC.1979.11327.

[17] Mithlesh Arya, Namita Mittal, and Girdhari Singh. Cervical cancer detection using single cell and multiple cell histopathology images. In *Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics*, pages 205–215. Springer Singapore, 2019. doi: 10.1007/978-981-13-8300-7_17. URL https://doi.org/10.1007/978-981-13-8300-7_17.

[18] S. Jaya and M. Latha. Channel based threshold segmentation of multi-class cervical cancer using mean and standard deviation on pap smear images. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pages 721–726, 2020. doi: 10.1109/ICESC48915.2020.9156020.

[19] Jata S Misra, Vinita Das, Anand N Srivastava, Uma Singh, and Madhulika Singh. Agnor counts in cervical smears under normal and other cytopathologic conditions. *Analytical and Quantitative Cytology and Histology*, 27(6):337–340, 2005.

[20] Madhulika Singh, Sahdeo Prasad, Neetu Kalra, Uma Singh, and Yogeshwer Shukla. Silver-stained nucleolar organizer regions in normal and dysplastic cervical lesions: Correlation with DNA ploidy and s-phase fraction by flow cytometry. *Oncology*, 71(5-6):411–416, 2006. doi: 10.1159/000107773. URL https://doi.org/10.1159/000107773.

[21] Jata S. Misra, Chhavi, Madhulika Singh, Anand N. Srivastava, and Vinita Das. Assessment of potential of AgNOR counts as tumor marker in cervical carcinogenesis. *Diagnostic Cytopathology*, 36(3):194–195, 2008. doi: 10.1002/dc.20718. URL https://doi.org/10.1002/dc.20718.

[22] Carolina Marian Pedrini et al. Análise das regiões nucleolares com prata (agnor) em lesões cervicais. 2020. URL https://repositorio.ufsc.br/xmlui/handle/123456789/216317.

[23] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
[24] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *NIPS-W*, 2017.
[25] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. https://github.com/facebookresearch/detectron2, 2019.
[26] David MW Powers. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*, 2020.
[27] Tom Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27 (8):861–874, 2006. ISSN 0167-8655. doi: https://doi.org/10.1016/j.patrec.2005.10.010. URL https://www.sciencedirect.com/science/article/pii/S016786550500303X. ROC Analysis in Pattern Recognition.