

# What is the State of the Art of Computer Vision-Assisted Cytology? A Systematic Literature Review

André Victória Matias<sup>a,\*</sup>, João Gustavo Atkinson Amorim<sup>a</sup>, Luiz Antonio Buschetto Macarini<sup>b</sup>, Allan Cerentini<sup>a</sup>, Alexandre Sherlley Casimiro Onofre<sup>d</sup>, Fabiana Botelho De Miranda Onofre<sup>d</sup>, Felipe Perozzo Daltoé<sup>c</sup>, Marcelo Ricardo Stemmer<sup>b</sup>, Aldo von Wangenheim<sup>e</sup>

<sup>a</sup>*Department of Informatics and Statistics, Federal University of Santa Catarina,  
Florianópolis, Brazil*

<sup>b</sup>*Automation and Systems Department, Federal University of Santa Catarina,  
Florianópolis, Brazil*

<sup>c</sup>*Department of Pathology, Federal University of Santa Catarina, Florianópolis, Brazil*

<sup>d</sup>*Clinical Analyses Department, Federal University of Santa Catarina, Florianópolis,  
Brazil*

<sup>e</sup>*Brazilian Institute for Digital Convergence, Federal University of Santa Catarina,  
Florianópolis, Brazil*

---

## Abstract

Cytology is a low-cost and non-invasive diagnostic procedure employed to support the diagnosis of a broad range of pathologies. Cells are harvested from tissues by aspiration or scraping, and it is still predominantly performed manually by medical or laboratory professionals extensively trained for this purpose. It is a time-consuming and repetitive process where many diagnostic criteria are subjective and vulnerable to human interpretation. Computer Vision technologies, by automatically generating quantitative and objective

---

\*Corresponding author

Email addresses: [andre.v.matias@posgrad.ufsc.br](mailto:andre.v.matias@posgrad.ufsc.br) (André Victória Matias ), [joao.atkinson@posgrad.ufsc.br](mailto:joao.atkinson@posgrad.ufsc.br) (João Gustavo Atkinson Amorim), [luiz.buschetto@posgrad.ufsc.br](mailto:luiz.buschetto@posgrad.ufsc.br) (Luiz Antonio Buschetto Macarini), [allan.c@posgrad.ufsc.br](mailto:allan.c@posgrad.ufsc.br) (Allan Cerentini), [alexandre.onofre@ufsc.br](mailto:alexandre.onofre@ufsc.br) (Alexandre Sherlley Casimiro Onofre), [fabiana.onofre@ufsc.br](mailto:fabiana.onofre@ufsc.br) (Fabiana Botelho De Miranda Onofre), [felipe.daltoe@ufsc.br](mailto:felipe.daltoe@ufsc.br) (Felipe Perozzo Daltoé), [marcelo.stemmer@ufsc.br](mailto:marcelo.stemmer@ufsc.br) (Marcelo Ricardo Stemmer), [aldo.vw@ufsc.br](mailtoaldo.vw@ufsc.br) (Aldo von Wangenheim)

descriptions of examinations' contents, can help minimize the chances of misdiagnoses and shorten the time required for analysis. To identify the state-of-art of computer vision techniques currently applied to cytology, we conducted a Systematic Literature Review, searching for approaches for the segmentation, detection, quantification, and classification of cells and organelles using computer vision on cytology slides. We analyzed papers published in the last 4 years. The initial search was executed in September 2020 and resulted in 431 articles. After applying the inclusion/exclusion criteria, 157 papers remained, which we analyzed to build a picture of the tendencies and problems present in this research area, highlighting the computer vision methods, staining techniques, evaluation metrics, and the availability of the used datasets and computer code. As a result, we identified that the most used methods in the analyzed works are deep learning-based (70 papers), while fewer works employ classic computer vision only (101 papers). The most recurrent metric used for classification and object detection was the accuracy (33 papers and 5 papers), while for segmentation it was the Dice Similarity Coefficient (38 papers). Regarding staining techniques, Papanicolaou was the most employed one (130 papers), followed by H&E (20 papers) and Feulgen (5 papers). Twelve of the datasets used in the papers are publicly available, with the DTU/Herlev dataset being the most used one. We conclude that there still is a lack of high-quality datasets for many types of stains and most of the works are not mature enough to be applied in a daily clinical diagnostic routine. We also identified a growing tendency towards adopting deep learning-based approaches as the methods of choice.

*Keywords:* Cytology, Segmentation, Classification, Deep Learning, Computer Vision

---

## 1. Introduction

Cytology is a diagnostic technique where cells are harvested from tissues by aspiration or scraping, prepared with different staining techniques and examined on the microscope. Since first described by Dr. Alfred François Donné in 1837 when he discovered *Trichomonas vaginalis*, this technique has been used to support the diagnosis of a broad range of pathologies. It became popular in 1941 when Dr. Georgios Papanicolaou proved cytology was useful in the diagnostic of malignant cells of vaginal smears of the uterus [1].

Cytology has been successfully applied to the diagnostic of some of the

most common and deadly cancers of the human body such as those from bone marrow, lungs, pancreas, breast, cervix, thyroid, oral cavity, and many others. Cytology has also been useful to diagnose a complex variety of non-cancer related pathologies such as inflammatory conditions and also pathogens as fungus, bacteria, and virus [2]. Collecting cells by aspiration or scraping methods presents the advantage of having a lower cost and being less invasive and easier to be performed by clinicians when compared to other methods of harvesting tissue samples for diagnosis, such as the tissue biopsy [3].

In cytology, the pathological status of the cells that have been collected can be analyzed by different approaches, such as DNA quantification, identification of cell surface key proteins, presence of pathogens, and, most often, cell morphology, describing the shape and size of cells and their organelles. Each method requires different cell staining and/or processing to highlight specific organelles or parts of the cell, but regardless of the method, the final analysis is usually performed manually under a light microscope [2]. Although this procedure is performed by medical professionals extensively trained for that, many of the diagnostic criteria are therefore vulnerable to human interpretation and bias [4]. Hence, more recent diagnostic approaches are associated with different diagnostic methods and computer technologies in order to reduce the chances of misdiagnoses. Another reason to employ computational methods is the long processing time that a manual analysis process affords [5]. Computer-based methods have the potential of not only increasing the degree of objectiveness and reproducibility of analyses but also their speed [6].

As the algorithms become more robust and the image analysis techniques more powerful, more information from cytology samples can be collected in an automated way [7].

Machine Learning, especially Artificial Neural Networks (ANNs), have been employed, with varying results, to develop Computer Vision (CV) methods for the automated support of quantitative analysis and diagnostic of cytological samples for over 25 years now, with methods ranging from spatial information parameter extraction with graphs and posterior ANN classification [8, 9], comparison between ANNs and multivariate statistical methods [10, 11], and comparison between simple ANNs, self-organizing semantic maps and especially developed nearest-neighbor-based ANN models [12].

In the last years, however, ANN-based Deep Learning (DL) methods have revolutionized the field of Neural Networks [13]. This revolution occurred mainly in image and signal analysis tasks, and achieved remarkable perfor-

mance in different computational tasks, showing its robustness and effectiveness on feature extraction from data in different scenarios [14]. Recent works using DL approaches in genomics and biomedical applications demonstrated the flexibility of this approach in handling complex problems, turning DL methods into a very promising approach for automatically analysing cytometry data [5].

On the other side, several factors can potentially introduce sample acquisition variations, such as microscope particularities, the quality of the staining chemicals, and the lab process. Considering this, to provide reliable cytology CV-based analysis methods, it is necessary to develop algorithms that are robust to this kind of variation. Figure 1 shows some examples of cytology image fields acquired in a Brazilian laboratory to illustrate such variations.

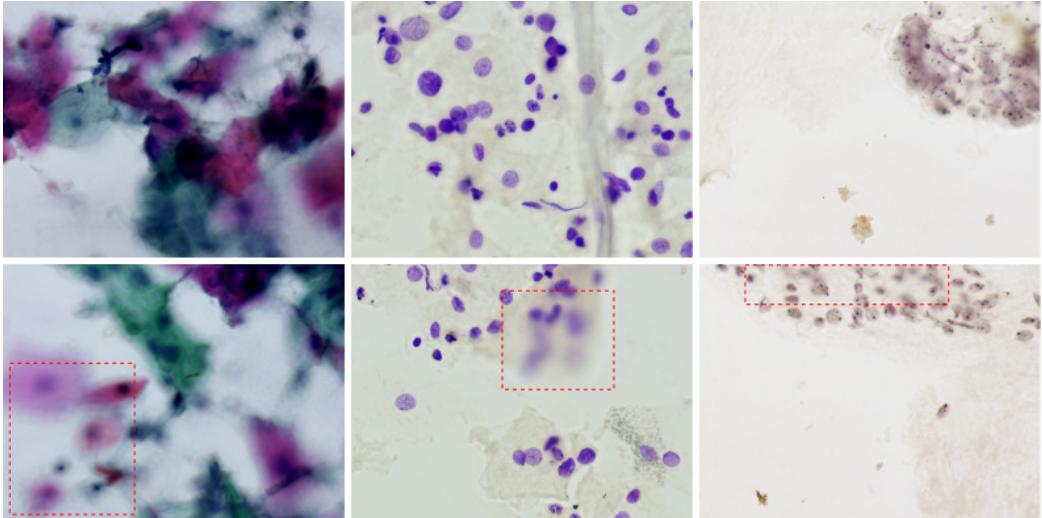


Figure 1: Examples of problems that can be found on the clinical routine. In the first row, highly overlapping cells. On the second row, inside the red-dotted rectangles, unfocused parts of the image fields.

Another scenario that reliable CV methods for cytology could make possible is that of *large-scale tele-cytology*: personnel in labs in small cities in upstate locations are trained to perform the examination collection, samples staining, and image acquisition through digitizing. These samples are then run through image processing routines to perform the quantitative analyses and feature extractions that can be automated. These CV analyses can

be performed either on the cloud or local computers, but the software tools will be able to be operated by local upstate personnel. Later, automatically generated reports from these analyses, together with the original digitized slides, can be uploaded into a telemedicine system, where a specialist can review the images and the quantitative data and provide a final findings report. This method would enable a more distributed, much faster, and less specialist-dependent cytology procedure, concentrating the workforce of experts on providing the final analyses and conclusive findings reports. However, to implement such a scenario, it is necessary to develop more robust CV techniques that account for larger quality variations on sample preparation and image acquisition. This scenario means that mature, robust, and clinically applicable CV methods in cytology could not only provide better means to perform quantitative cytology but could also represent a paradigm change in a diagnostic process that, even in a digital era, in many places still depends largely on the physical transportation of samples. To inquire into the state-of-the-art CV methods for cytology, we performed a systematic review of the literature (SLR).

## 2. Materials and Methods

A SLR is conducted based on the systematic review protocols for the Computer Sciences field elaborated by Kitchenham [15]. In this review, we searched for approaches for cell segmentation, detection, and classification using computer vision on cytology slides images. Analysing papers published in the last 5 years (from 2016 to 2020), we evaluated the tendencies and main problems present in this research area, highlighting the computer vision methods, staining techniques, result evaluation metrics, and the availability of the used datasets. Our review structure definition is shown in Figure 2. In this review, we focused on identifying commonalities, differences, and tendencies among the different works described in the literature. We provide concise, graphic, and tabular representations of these characteristics and relationships and discuss them. We also published a complementary technical report [16] where we focused on analysing some of the papers individually and in more detail.

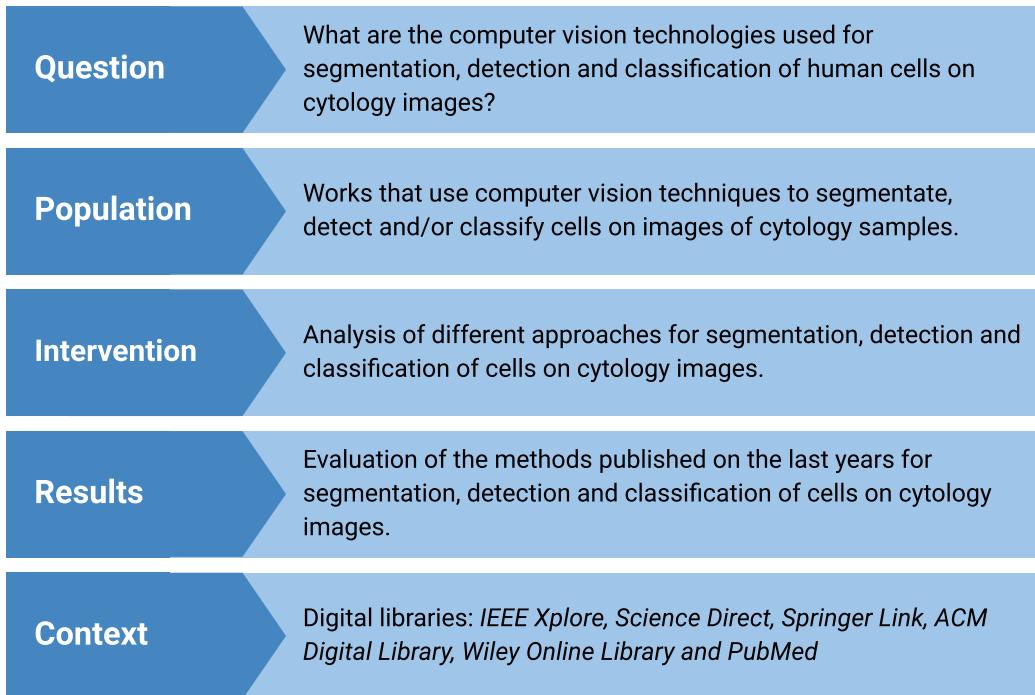


Figure 2: Review definition

### 2.1. Search Definitions

The details of the search locations (databases), terms, inclusion and exclusion criteria are presented in Tables 2, 3, 4, 5, 6, and 7. The Papanicolaou is the only cytology technique explicitly used in the search terms because a considerable amount of authors does not cite it as a cytology or cytopathology examination and relevant papers would not have been included if these more specific terms were not used.

To avoid *evidence selection bias*, [17] we did not limit our search only to metasearch databases such as *PubMed* or *NCBI*, which tend to present only results from references previously classified as belonging to health sciences and biomedical topics. Instead, we additionally conducted our SLR as it is commonly performed in the field of Computer Sciences (CS), also accessing directly the databases of the individual scientific publishers. We did this because this SLR is focused on *CS research on technology applied to cytology* and a requirement was that relevant CS technology journals and conferences should also be fully taken into consideration. Table 2 presents the IEEEExplore Search Definitions.

Search Locations	Search Terms
<ul style="list-style-type: none"> <li>- IEEE Xplore;</li> <li>- Science Direct;</li> <li>- Springer Link;</li> <li>- ACM Digital Library;</li> <li>- Wiley Online Library;</li> <li>- PubMed.</li> </ul>	<ul style="list-style-type: none"> <li>- “cytology” OR “cytopathology” OR “Papanicolaou” OR “Pap smear”;</li> <li>- “image” OR “vision”;</li> <li>- “classification” OR “segmentation” OR “detection”;</li> </ul>
<b>Inclusion Criteria</b>	
<ul style="list-style-type: none"> <li>- Papers that use computer vision to segment, detect and/or classify cells on cytology images;</li> <li>- Studies performed in humans.</li> </ul>	
<b>Exclusion Criteria</b>	
<ul style="list-style-type: none"> <li>- Papers that do not use computer vision techniques to analyse cytology images;</li> <li>- Studies not performed on human cells;</li> <li>- Papers that focus on image enhancement;</li> <li>- Short papers, search definition, supers, such as abstracts or expanded abstracts;</li> <li>- Review articles;</li> <li>- Articles that were not completely available on the database.</li> <li>- Papers not written in English.</li> </ul>	

Table 1: Search definition

## 2.2. Search Execution

The initial search was made on September 16th, 2020 and resulted in 431 papers which were exported to Rayyan QCRI<sup>1</sup> where 13 duplicates were removed and the remaining ones were divided between four reviewers that read the titles and abstracts of each paper and filtered them based on the inclusion/exclusion criteria. After reading the full texts, the reviewers selected 157 papers in total.

---

<sup>1</sup><https://rayyan.qcri.org/>

<b>Database</b>	<b>IEEEXplore</b>
<b>Date</b>	2016 to 2020
<b>Filters</b>	Any Formats
<b>Number of Works</b>	89
<b>Selected Works</b>	68
<b>Search String</b>	(((((“Abstract”:cytology OR “Abstract”:cytopathology OR “Abstract”:Papanicolaou OR “Abstract”：“Pap smear”) AND (“Abstract”:image OR “Abstract”:vision) AND (“Abstract”:classification OR “Abstract”:segmentation OR “Abstract”:detection)) OR ((“Document Title”:cytology OR “Document Title”:cytopathology OR “Document Title”:Papanicolaou OR “Document Title”：“Pap smear”) AND (“Document Title”:image OR “Document Title”:vision) AND (“Document Title”:classification OR “Document Title”:segmentation OR “Document Title”:detection)) OR ((“Author Keywords”:cytology OR “Author Keywords”:cytopathology OR “Author Keywords”:Papanicolaou OR “Author Keywords”：“Pap smear”) AND (“Author Keywords”:image OR “Author Keywords”:vision) AND (“Author Keywords”:classification OR “Author Keywords”:segmentation OR “Author Keywords”:detection))))

Table 2: IEEEXplore Search Definitions

Next, Table 3 presents the Science Direct Search Definitions.

<b>Database</b>	<b>Science Direct</b>
<b>Date</b>	2016 to 2020
<b>Filters</b>	Any Formats
<b>Number of Works</b>	135
<b>Selected Works</b>	30
<b>Search String</b>	((cytology OR cytopathology OR Papanicolaou OR ”Pap smear”) AND (image OR vision) AND (classification OR segmentation OR detection)))

Table 3: Science Direct Search Definitions

In Table 4 is presented the Springer Link search definitions.

<b>Database</b>	<b>Springer Link</b>
<b>Date</b>	2016 to 2020
<b>Filters</b>	Any Formats; Discipline: Computer Science
<b>Number of Works</b>	115
<b>Selected Works</b>	28
<b>Search String</b>	(“cytology” OR “cytopathology”) OR (“Papanicolaou” OR “Pap smear”) AND (“image” OR “vision”) AND (“classification” OR “segmentation”) OR “detection”)

Table 4: Springer Link Search Definitions

The ACM Digital Library Search Definitions are shown in Table 5.

Database	ACM Digital Library
Date	2016 to 2020
Filters	Any Formats
Number of Works	15
Selected Works	10
Search String	[[[Abstract: cytology] OR [Abstract: cytopathology] OR [Abstract: papanicolaou] OR [Abstract: “pap smear”]] AND [[Abstract: image] OR [Abstract: vision]] AND [[Abstract: classification] OR [Abstract: segmentation] OR [Abstract: detection]]] OR [[[Publication Title: cytology] OR [Publication Title: cytopathology] OR [Publication Title: papanicolaou] OR [Publication Title: “pap smear”]] AND [[Publication Title: image] OR [Publication Title: vision]] AND [[Publication Title: classification] OR [Publication Title: segmentation] OR [Publication Title: detection]]] OR [[[Keywords: cytology] OR [Keywords: cytopathology] OR [Keywords: papanicolaou] OR [Keywords: “pap smear”]] AND [[Keywords: image] OR [Keywords: vision]] AND [[Keywords: classification] OR [Keywords: segmentation] OR [Keywords: detection]]] AND [Publication Date: (01/01/2016 TO 12/31/2020)]]

Table 5: ACM Digital Library Search Definitions

In Table 6 the Wiley Online Library search definitions are shown.

Database	Wiley Online Library
Date	2016 to 2020
Filters	Any Formats
Number of Works	45
Selected Works	7
Search String	[[[Keywords: ‘cytology’] OR [Keywords: ‘cytopathology’] OR [Keywords: ‘pap smear’] OR [Keywords: ‘papanicolaou’]] AND [[All: ‘computer vision’] OR [All: ‘image analysis’]] AND [[All: ‘classification’] OR [All: ‘detection’]] OR [All: ‘segmentation’]] AND [Earliest: (01/01/2016 TO 03/31/2020)]]

Table 6: Wiley Online Search Definitions

Next, in Table 7, the PubMed Search Definitions are presented.

Database	PubMed
Date	2016 to 2020
Filters	Any Formats
Number of Works	32
Selected Works	14
Search String	(“Papanicolaou” OR “Pap smear” OR “conventional cytology” OR “liquid-based cytology” OR “brush biopsy” OR “brush cytology”) AND (“Convolutional Neural Networks” OR “Computer Vision” OR “Deep Learning” OR “Image Processing”) AND (“classification” OR “segmentation” OR “detection”)

Table 7: PubMed Search Definitions

### 3. Analysis of the Papers

In this section, we present the analysis of the papers included in the review. Section 3.3 presents the staining techniques used in the analyzed papers along with a short explanation of each one and Section 3.4 describes the public datasets we found while reading the papers. In Section 3.5, we explain the different techniques used for analysis of cytology images in the analyzed papers, grouping them by method (Image Classification, Object Detection, and Segmentation) and by approach (Classic, Deep Learning, and Hybrid).

#### 3.1. Related Reviews

Similar works describing systematic reviews on cytology were found during the analysis of the article. Those works were excluded from our SLR for not presenting a method. It is, however, important to cite and analyse them: As shown in Table 8, these review papers focus on specific anatomies, specifically *breast* and *cervix*, whereas our SLR addresses Cytology in general. Another important aspect to be noticed is that most of these works focus on specific methods such as ML or DL, whereas we considered both classic CV and machine learning-based methods. Taking this into consideration, we understand that this work has a much larger scope than the related works

Paper	Year	Quantity	Body part	Method
[18]	2002-2017	30	breast	Machine Learning
[19]	until Feb, 2021	33	cervix	Deep Learning
[20]	1977 - 2018	78	cervix	Segmentation
[21]	1965 - 2016	137	breast	Any
Ours	2016-2020	157	Any	Any

Table 8: Related Reviews

and also includes the review of more recent articles than most of the cited works.

### 3.2. Classification Rationale for Computer Vision Approaches

In this paper we roughly divide the analyzed papers, from the point of view of the image processing technology used, into three general categories: (a) classic CV approaches (CCV), (b) deep learning (DL) approaches, and (c) hybrid CV approaches (HCV), where DL and CCV are employed together.

The general principle behind the CCV paradigm is that an image interpretation task starts from meaningless pixels and moves stepwise towards a meaningful representation of that image’s content. The transform steps move towards a growing abstraction and simplification of the image to find and isolate the elements of the image that have a meaning for a given particular application context. This is performed as a pipeline of transforms in different *levels of abstraction*, whereas the first transforms in this pipeline are transforms from images into new and simplified images ( $I \rightarrow I$ ), such as noise filters, border detectors or segmentation algorithms. Later, when the abstraction level grows, these transforms are performed from images into models ( $I \rightarrow M$ ) that represent descriptions of specific elements or the content of these images, such as segment color and shape parameters or texture descriptors. Further on, these models are transformed into other, more abstract models ( $M \rightarrow M$ ) that describe the meaning or classification of those objects [22]. These transform steps can be applied on three different *image processing domains*: (i) *value*, when only context-less individual pixel values are being considered, such as in histograms or thresholds; (ii) *space*, when pixels are considered spatially, in the context of their neighborhoods, such as in border detectors, segmentation methods or general convolutions; and (iii) *frequency*, when pixel variation patterns along the image are taken into consideration, such as in Fourier or Wavelet transforms [22].

One characteristic of CCV is that these pipelines are extremely problem- and image type-specific and strongly parameter-dependent. This means that a processing pipeline that works well for a given type of image content, e.g., Papanicolaou samples, will not work for Feulgen-stained samples. This parameter sensitivity also means that, e.g., a segmentation algorithm that works well for a set of images, will not work as well with another set of images of the same kind, but that shows different lightness and contrast characteristics. This made CCV solutions challenging to develop, requiring deep mathematical knowledge of the individual methods in order to allow an adequate composition of a pipeline for a specific problem. These characteristics also made CCV solutions less robust and highly image-quality-dependent, which has led to some image processing cytology solutions to require the user to perform a rigorous process for the preparing and digitizing of the samples, to guarantee that the image processing pipeline actually works [23].

The advent of DL techniques broke the CCV paradigm: DL convolutional neural networks (CNN) can learn sequences of convolution operations that represent image transforms in all three image domains and all levels of abstraction [14]. Due to their deep sequential structure, DL CNNs can also learn hierarchical transformation and representation sequences [14]. A DL application is able to, in one single step, start from pixels and end with a representation of the meaning of a given image. DL also simplified the development process of CV applications: the main steps of a CV solution can now be *trained*. However, DL solutions have the disadvantage of being black-box processes: the convolution operations performed by the network are learned and coded as data in the structure of a given network, and not separated processes that can be isolated and individually analyzed.

An overview of both paradigms is shown in Figure 3. For this work, we considered CCV approaches all those that did not use any DL technique in its processing pipeline.

### 3.3. Staining Techniques

Cytology examinations are performed through the use of microscopes to evaluate slides prepared from small sets of cells. Usually, a staining technique must be used to highlight the cells' desired features. Figure 4 shows part of an unstained slide and Figure 5 shows an example of the results of applying different staining and highlighting techniques that we have identified in the analyzed papers. The techniques applied in these papers are Papanicolaou (130 papers), H&E (20 papers), Giemsa (5 papers), Feulgen

		Step/Action in the Processing Pipeline			
Paradigm	Domain	I. Filtering and Pre-Processing	II. Conditioning and Simplifying	III. Labelling and Description	IV. Interpretation and Description
	Enhance the image	Simplify an image	Simplify an image	Describe an image	Reason about an image
	1. Value	Windowing and scale transform	Thresholding, arithmetic and logic operations with pixels	Histogram analysis	
Classic Computer Vision	2. Space	Despeckling, mathematical morphology operations, convolutional and diffusion filters	Convolutional-based edge detection (Marr, Canny, Sobel, etc), saliency detection (Shaa Shua, etc), region-growing segmentation (watershed, Mumford-Shah, CSC, etc), skeletonization, split & merge	Spatial classifiers, classic ANNs (backpropagation, Boltzmann, Kohonen), statistic classifiers, classic machine learning (SVM, IBL, random forests, etc), rule-based classifiers, feature detection (SIFT, SURF, etc)	Spatial data structures (octrees, quadtrees), face-based models (DLFL, etc), boundary models, graph-based and spatial grammars, semantic networks
	3. Frequency	FFT- and wavelet-based filters (low-, band- and high-pass, etc)	FFT- and wavelet-based operations for frequency elimination, grouping and separation	FFT- and wavelet-based classification methods, Gabor filters	
Deep Learning	Convolutional Neural Networks	AlexNet, VGG, ResNet, Inception, YOLO, SSD, Mask-RCNN, etc			
	Transform/Mapping	$I \rightarrow I$	$I \rightarrow I$	$I \rightarrow M$	$M \rightarrow M$

Figure 3: Computer vision domains, steps and paradigms.

(5 papers), Hematoxylin+DAB (1 paper), AgNOR (1 paper), and 1 paper that employed unstained slides. There are also 3 papers that do not specify clearly which staining method was used.

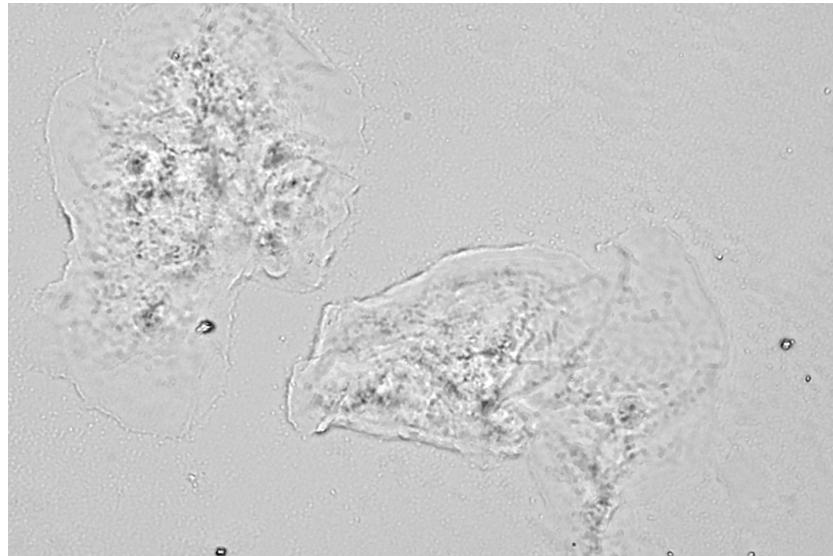


Figure 4: Example of an unstained cytology slide

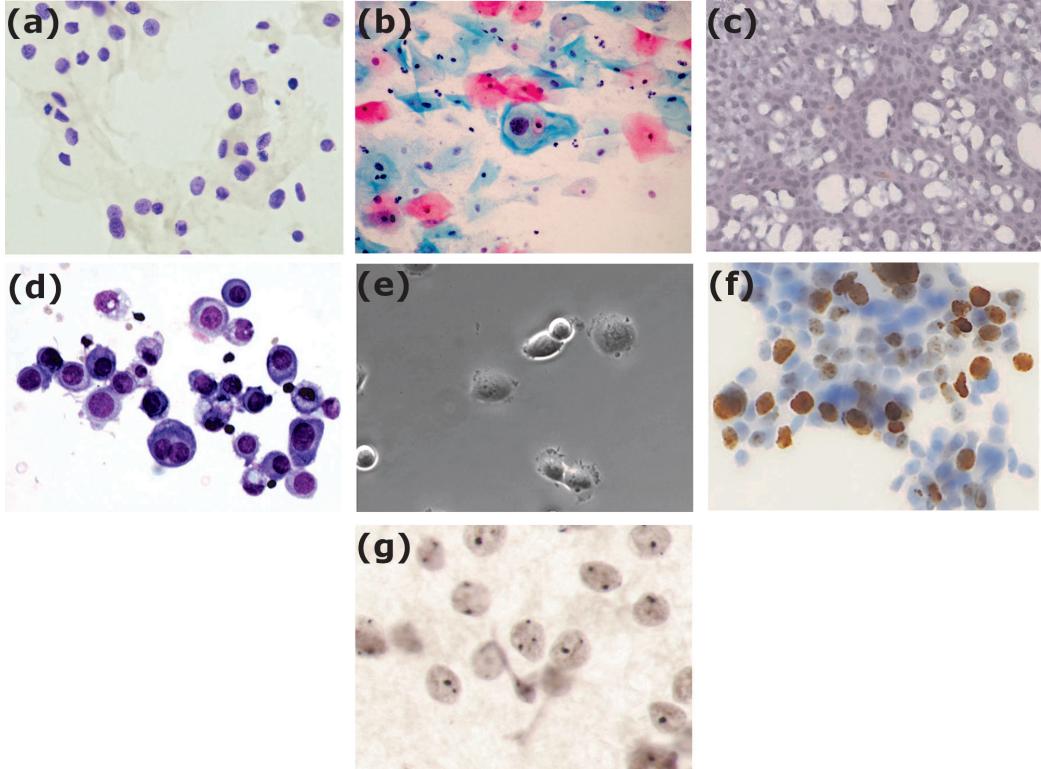


Figure 5: Examples of staining techniques. (a) Feulgen, (b) Papanicolaou, (c) H&E [24], (d) Giemsa, (e) Unstained using phase-contrast microscopy [25], (f) Hematoxylin+DAB [26], (g) AgNOR

Except for [25], research described in all papers employed some kind of staining technique, with Papanicolaou being the most frequently used. Papanicolaou is an inexpensive staining technique, generally applied to detect cervical cancer, composed of five dyes (Hematoxylin, Orange G, Bismark brown Y, light green SF yellowish, and eosin Y) that aims to define nuclear details, highlight cell boundaries, and differentiate between types of cells. Although it is widely used, the Papanicolaou staining is not a well-standardized method, which produces a considerable range of variations in its results [27]. Hematoxylin is a dye that stains the chromatin and is generally used along with eosin (H&E), a dye that stains the nucleoli and the cytoplasm of the cells [27]. [28] states that Raman spectroscopy is more effective than the use of H&E staining for feature extraction. One of the analyzed papers [26] used DAB (3,3'-Diaminobenzidine) instead of eosin for immunohistochemistry.

cal staining. The Giemsa stain is one of the Romanowsky-type stains, which allows the estimation of relative cell and nuclear sizes and the visualization of cytoplasmic details, smear background elements, and intercellular matrix components [27]. Feulgen staining is used to identify chromosomal material or DNA in the cells specimens [27] and the AgNOR technique consists of applying a silver nitrate staining to the slide to highlight Nucleolar Organizer Regions (NORs) [29].

### 3.4. Datasets

Data plays an essential role in machine learning, particularly for deep learning approaches [30]. When dealing with medical data segmentation, the data availability is often a problem [31], sometimes being necessary to collect and label more samples. Publicly available datasets are important tools for benchmarking, providing means for a fair comparison between the methods published in the literature.

We were able to identify 12 publicly available datasets: HEMLBC [32], ISBI 2014 [33], ISBI 2015 [34], DCCL [35], Herlev [36], SIPaKMeD [37], Perolehan Citra Kanker Serviks [38], AGMC-TU Pap-Smear [39], CCAgT: Images of Cervical Cells with AgNOR Stain Technique [40, 41], UFSC-OCPap [42], and 2 unnamed datasets proposed in [43] and [44]. Except for HEMLBC, which consists of Liquid-based H&E slides, and the CCAgT, which consists of AgNOR slides, all the other datasets use the Papanicolaou staining technique. The DTU/Herlev dataset was used in 38 works, being the most recurrent among all datasets.

Ten of the twelve public datasets found are composed of samples collected from the Cervix, except for the dataset presented in [43], which is composed of Fine-needle Aspiration Cytology (FNAC) Smears collected from the breast, and the UFSC-OCPap dataset, which is composed of oral smear samples.

### 3.5. Image Classification

Image classification is a fundamental task in computer vision. This task aims at categorizing images into one or several predefined classes. Image classification can be considered the basis for other computer vision tasks, e.g., localization, detection, and segmentation [45]. Even though it is a trivial task for human beings, it is challenging for an automated system [46]. In the past few years, the developers usually adopted a two-step image processing pipeline approach to solve this problem. First, a set of image type-specific

features were extracted from images using a sequence of one or more hand-crafted feature descriptors. Afterwards, these features were used to train a classifier [47]. However, there were many complications when using this approach, such as viewpoint-dependent object variability and the high intra-class variability of having many object types [46]. One of the main problems of this approach was the fact that the accuracy of the classification task was profoundly dependent on the design and robustness of the specific feature extraction algorithms employed [47]. Nowadays, deep learning models are being used for integrated feature extraction and transformation, as well as for pattern analysis and classification, integrating all steps of this image processing pipeline into a single large classification step performed by a deep learning neural network. It has been shown to overcome these challenges [45].

### 3.5.1. Classic Approaches

In most of the papers we retrieved, classic CV approaches are used to extract features of the images for later classification using the Artificial Neural Network (ANN), Support Vector Machine (SVM) or K-Nearest Neighbors (kNN) techniques. In [48], however, the authors employ an approach based only on the contours of the nucleus, where the classifier uses the geometric features and the Fast Fourier Transform (FFT) to describe this external contour.

For the feature extraction step, multiple approaches were employed: Histogram analysis, Gray level Co-occurrence matrix (GLCM), Local binary pattern (LBP), Laws texture energy-based, discrete wavelet transform (DWT), Markov Random Fields, morphometric, textural and intensity. In [49], the authors used Scale-invariant Feature Transform (SIFT) for the feature extraction step. Classification using SVM was used in [50, 51, 52, 53, 54, 55, 56, 57, 49, 58, 59], ANN in [60, 51, 53, 61, 62, 54, 63, 64, 65], Decision Tree with Logistic Regression in [66] and Random Forest (RF) in [23, 54, 58]. In [67], the authors used kNN, SVM, and ANN to extract features from manually annotated cytoplasm and nuclei. In [39], the authors published their dataset (AGMC-TU Pap-Smear), where they perform a benchmark among models of ANN, SVM, kNN, Linear Discriminant Analysis (LDA), RF, decision tree and Naive Bayes to classify the pap-smears cells in normal and abnormal. Among these papers, 2 work with oral cavity samples, 7 use cervical Pap smear samples, and in [23] the authors use lung cells slides stained with Feulgen. In [68], the authors propose using the Randomized Neural Net-

work Texture Signature method to classify cells in Pap smear slides. Global Significant Value, statistical-based texture features and interquartile range were used in [65] to extract cell features used as input to a Naive Bayes model, a J48 tree and an ANN for classification. In [69], cell features were classified into normal and abnormal using two classifiers: a two-class boosted decision tree and a two-class logistic regression model. The best result between the classifiers was passed to multi-class logistic regression classifiers to classify the abnormal cells into 4 classes. In [70], the authors compare an ANN, combined with a genetic algorithm for feature optimization, and a Hierarchical Decision Approach (HDA). The results indicate that the HDA had the best results. The work from [58] uses the 2-D Fourier transform to extract spectral features that are used to train the following classical machine learning algorithms: kNN, SVM, Random Forest, and AdaBoost.

### *3.5.2. Deep Learning Approaches*

The Deep Learning (DL) approaches we retrieved employ the most recent deep convolutional neural network models seen in the state-of-the-art CV research, such as Residual Neural Network (ResNet) models [71] and VGG net [72].

In this sub-category, we were able to identify research works employing five different stains and cells from four different anatomies. In [25], the authors used a Convolutional Neural Network (CNN) model to select single cell images in the first stage, and in the second stage, a CNN based on *VGG\_CNN\_F* [73] to classify human bladder cancer cells was used. In [74], the authors proposed a novel deep neural network to classify microscopy images. They concatenate the features extracted using three pre-trained deep CNNs. These features are used to train two fully-connected layers to perform classification.

In [75], the authors have used cervical slides stained with Feulgen to realize a manual feature extraction and classify the cells into 10 categories using the InceptionV3 [76] model. The only DL classification work that used images of nasal mucosa cells with the May Grunwald-Giemsa (MGG) stain was [77]. They have used a small CNN with two layers proposed to classify the cells into seven distinct classes. We also identified two papers that used the H&E-stain. In [78], the authors trained four models: VGG-16, VGG-19, ResNet-50 and ResNet-101. The ResNet-101 was the model described as being the one that showed the best overall performance.

In [79], the authors also employed H&E staining with breast FNAC sam-

ples, which were classified using the GoogLeNet [80] model. In this paper, the authors propose a voting pre-processing to select the fields of the slide used as input in order to classify it. In [81], the authors developed an algorithm to evaluate thyroid fine-needle aspiration biopsy using whole slide images. The designed algorithm is based upon two CNNs: one to identify the follicular groups and another to predict the Bethesda System for Reporting Thyroid Cytopathology category and the final pathology. Both networks are based on the VGG-11 model, pre-trained on the ImageNet dataset. According to the authors, the performance of the algorithm was comparable to that of an expert cytopathologist. In [82], the authors proposed a system that displays regions of interest on whole slide images containing potential anomaly. Then, the pathologists review these regions to increase productivity. The authors built multiple deep learning classifiers, using VGG-16 and DenseNet-121, as part of the proposed architecture.

We analyzed more than ten papers that employed the Papanicolaou staining technique. To classify lung cells, the authors used a model based on VGG-16 to classify pulmonary images in [83]. [84] compared the ability of AlexNet, GoogLeNet, ResNet, Inception V3, VGG and DenseNet to classify the cells into eleven classes. In [85], the authors compared AlexNet, GoogLeNet, ResNet, and DenseNet to classify nuclei in the Herlev dataset. A Deep Belief Network made of Restricted Boltzmann Machines was proposed in [48], where the model was compared to a classic approach. There are 19 papers that showed similar approaches with different models to classify Pap Smear images, where [86, 35, 87, 88, 89, 90, 91, 42] used the ResNet model, [35, 88, 89, 90, 92] used the Inception model, [88, 89, 90] used the VGG model, [35, 90] used the DenseNet model, [93] used a combination of the ResNet model and the DenseNet model, [48] used a Deep Belief Network, [94, 95] used the AlexNet model, [88] used the SqueezeNet model, and [96, 97, 98, 99, 100, 101] used custom CNN models.

The approach used in [87] was divided into two steps: In the first, Pap Smear images were segmented into nuclei, cytoplasm and background, employing a U-net model using a ResNet as its backbone. In the second step, the nuclei were classified into sub-types. Combining a CNN with multiple instance learning, the authors in [96] proposed a model for dysplasia detection. The authors of [88] and [102] propose a two-step method where a CNN is used to extract the image features that are then passed to a classifier to produce the final result. In [88], they compare VGG-19, SqueezeNet, Resnet-50, and InceptionV3 to extract the images features and kNN, RF

and SVM for classification. In [35], the models ResNet-101, InceptionV3, and DenseNet-121 were compared to classify cells in Pap Smear images. In [93], for pre-cancer lesions classification, the models AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-101, and GoogLeNet were compared. The work from [103] uses a network composed of 10 convolutional layers and three fully connected layers pre-trained on ImageNet. Then, the proposed method discards all the non-relevant kernels according to a score and the network is trained using a private cervical pap smear dataset.

### 3.5.3. Hybrid Approaches

We considered as Hybrid approaches the ones that employed classical image processing/computer vision methods for pre-, post-processing or image enhancing, for example, together with a DL step for image classification. In [43], the authors considered the cytoplasm and red blood cells present in the images as noise. Then, Histogram equalization was employed in the red channel output as a pre-processing step to remove the cytoplasm. Afterwards, images were segmented to remove the cytoplasm and the red blood cells. The global and the Otsu thresholding were employed to make it simple. According to the authors' visual investigation, Otsu thresholding performed better than global thresholding. This method was used to isolate the nuclei in each image, and then the authors compared four different CNN models for the classification of FNAC samples.

In [104], the authors proposed a three-layer CNN to classify nasal cytology samples. In the pre-processing step, aiming to create a system capable of recognizing images of slides prepared with different techniques, the Otsu algorithm was applied in grayscale images. The morphological dilation and the Watershed algorithm were applied. Then, the next step is labelling, for marking the different objects with different shades of colour to facilitate subsequent classification. In [37], the authors presented a new database of cervical images where the cells were categorized into 5 classes. Among other classic methods, the authors employed a VGG-19 for cell classification. The work from [105] extracts features from the shape of the cell and use additional methods as Local Binary Pattern (LBP), Gray Level Co-Occurrence Matrix (GLCM) and Shape Features to extract additional features which were later used to train multiple Deep Belief Networks that were developed using multiple combinations of hidden layers and hidden units to perform the experiments and find out which is the best configuration.

### *3.6. Object Detection*

Object detection is a computer vision task that aims to localize each instance of certain object classes in an image. It can be seen as a combination of two concepts: the first, image classification, which consists of predicting the class of one object in an image; the second, object localization, refers to identifying the location of one or more objects in an image. Object detection combines these two tasks to localize and classify one or more objects in an image. Since the advent of CNNs, object detection has experienced impressive progress in terms of accuracy and speed.

Concerning object detection applied to cytology, we could identify that all the papers employed directly state-of-the-art object detectors. None of the authors proposed a new architecture of CNN. Among the different architectures used, two works used the YOLOv3 model. In [75], the authors used YOLOv3 [106] to detect cells in images collected from the cervix. A task-specific cascade classifier was used to improve the classification performance of hard examples. In [107], the YOLOv3 was used to detect cell nuclei on pleural effusion cytopathology images stained with Papanicolaou. According to the authors, the most significant contribution of the method was the detection speed.

In [35], it was presented, according to the authors, the largest cervical cytology dataset, called Deep Cervical Cytological Lesions (DCCL). The authors used a Faster R-CNN and a RetinaNet [108] with a ResNet-50 backbone as baseline detection models. For classification, the authors used three different CNN architectures: InceptionV3, ResNet-101, and DenseNet-121. In [109], a region detection and classification method based on multi-semantic label combined with morphological information analysis was proposed. The authors then used transfer learning to train a Faster R-CNN model, with VGG-16, ResNet50, and ResNet101 as backbones to classify the dataset.

The authors from [110] also used the Faster R-CNN, R-FCN, and SSD architectures for nuclei detection in pleural effusion pathology images. They used different feature extractors, namely, ResNet, Inception v2, and MobileNet for performance comparison. In [111], the authors proposed a framework to extract features from H&E stained images using pre-trained CNN architectures: GoogLeNet, VGG, and ResNet. The feature extractor output is then fed to a fully connected layer to classify malignant and benign cells using average pooling classification. The authors from [112] tested multiple detection network architectures applied to H&E and Papanicolaou stained slides to detect peritoneal metastasis in ascites cytopathology. In [113], the

authors used a Mask R-CNN model for nuclei detection in Pap smear slides and, in [91], an Fully Convolutional Regression Network (FCRN) is used to detect nuclei in cytology slide images that are after used to test ResNet50 and DenseNet201 for single-cell nuclei classification. The authors from [42] compared the Faster R-CNN and the RetinaNet models, using ResNet-50 as the backbone, to detect cell nuclei in Papanicolaou stained slides using both a binary and a multiclass approach.

Besides most of the detection papers analyzed use a CNN as the base model, in Mufidah et al. [114] the authors employed an Stacked Sparse Autoencoder (SSAE) along with a softmax classifier to detect nuclei in patches extracted using sliding windows. This method was tested to detect cell nuclei in Pap smear images containing overlapped cells.

### *3.7. Segmentation-Based Approaches*

Segmentation techniques in cytology aim to provide each pixel of an image with a label, classifying them between desired (e.g. cells or nuclei) and undesired objects (e.g. background), or between multiple classes (e.g. background, normal cell, abnormal cell, etc.), in order to interpret the image.

#### *3.7.1. Classic Approaches*

Classic segmentation approaches are based on operations over pixel intensity values, generally in the context of their neighborhoods, where the input usually is a matrix of color values that is processed to generate a matrix of the pixels' labels as the output. In the analyzed papers, this approach is used for detection or counting cells in 23 papers, segmentation of overlapping cells in 19 papers, and features extraction for classification in 40 papers. Due to the high dimensionality of the extracted features, 15 papers use a dimension reduction algorithm before passing the data to a classifier. Most of the selected features are based on the morphology and texture of the cells, although some papers use cell distribution information and [115] states that “moment-based features offer better separation between the malignancy classes than shape and texture based features”. The tools used for classification of the extracted features are SVM (20 papers), Relevance Vector Machine (RVM) [116], ANN (11 papers), kNN (6 papers), Naive Bayes (3 papers), Random Subspace-Linear Discriminant Analysis [117], k-means (2 papers), Fuzzy C-Means (5 papers), Histogram of Oriented Gradients (HOG) [118], Decision Tree (8 papers), Random Forest (9 papers), Logistic Regression (2 papers), LDA (3 papers), and Neighborhood Component Analysis (NCA) [117]. Also, in [119]

and [120], the authors use an ensemble classifier that combines the results of multiple algorithms to generate the final output using major voting.

To apply the segmentation processes, it is usually necessary to pre-process the image to highlight the desired features and smooth or eliminate undesired ones. The pre-processing techniques used in the selected papers were: conversion of the images' RGB color space to the grayscale (24 papers), HSV (5 papers), and  $l^*a^*b^*$  (7 papers) color spaces, besides noise reduction (32 papers), contrast adjustments (13 papers), convolutional filters (7 papers), gamma correction [121], and histogram equalization (7 papers). The pre-processed image is then passed to the actual segmentation algorithms, which are divided into three groups in this section: thresholding-based, trainable model-based, clustering-based and graph-based. Papers that use thresholding-based techniques (44 papers) apply a local or global threshold over the input image to divide its pixels between the background and the desired objects and to detect contours. Clustering-based techniques group the pixels based on a similarity measure using clustering algorithms like k-means (12 papers), Density-based spatial clustering of applications with noise (DBSCAN) (2 papers), Fuzzy C-means (4 papers), Simple Linear Iterative Clustering (SLIC) superpixel merging (6 papers), Mean shift [55], and Gaussian mixture model (3 papers). Graph-based approaches treat the image pixels and their relations as a graph and employ graph algorithms to segment the image, as shown in 6 papers. Trainable model-based methods use pixel-level information to train a model to distinguish between the regions of interest in the image and the background. These methods were used in 4 of the analyzed papers.

Moreover, these techniques usually rely on additional mathematical models, such as mathematical morphological operations (24 papers), active contours (7 papers), geometric shape constraints (17 papers), and other techniques (4 papers). These mathematical models are employed in the post and/or pre-processing phase to optimize the segmentation results. A table with details of each paper that apply classic segmentation-based methods is presented in Appendix B.

### 3.7.2. Deep Learning Approaches

In this section, we will describe papers that employed Semantic Segmentation and other CNN-based segmentation approaches. Semantic Segmentation (SS) is a CNN technique where each pixel of an input image is individually classified and, thus, the image is segmented into different parts and objects, grouping pixels that *semantically* belong together, i.e., belong to the same

class of the GT used to train the CNN [122]. The most widely used network architectures to solve segmentation problems are based on the U-Net [123] and SegNet [124] architectures. Both models use only convolutional layers and consist of *encoders* and *decoders*. The image that enters the network is transformed into a feature map and each encoder progressively reduces the size of these maps by condensing the features. The decoders do the opposite process: they take the output from the last encoder and progressively expand the feature maps, turning them into image segments.

Most DL papers [125, 24, 126, 127, 44, 87, 128, 129, 130, 131] applied classic neural networks like VGG, AlexNet, U-Net, Mask R-CNN, and ResNet to segment cells. Some of these papers applied extra steps to improve the segmentation accuracy. An SVM classifier was used in [24] to predict the nuclei presence probability in slide images. A voting ensemble method combining a fully convolutional network, a SegNet, and a U-Net applied to a CNN was used in [128]. A double U-Net, as known as W-Net, was used in [127].

Some papers used non-conventional neural networks with additional steps. The DeepLabv2 network with a modified U-Net, called TernausNet, was used in [132] and in [133] a modification of the U-Net is proposed, called PGU-net+, where a residual module is added to the model and it is trained using progressive growing input sizes. Cascade regression chain was used with a custom neural network in [134]. A binary tree-like network with two-path fusion attention failure was used in [135]. The authors from [136] implemented a Fully Convolutional Neural Network (FCN) model based on the U-Net. In [137], the authors presented the IRNet, a novel instance relation network for overlapping cell segmentation. The authors from [138] propose a model based on the U-Net with residual blocks to use as a framework to assist in the segmentation of nuclei, cells, and glands. They trained the model in histopathology datasets and validated it using a Pap smear dataset, showing that it can also segment cells in cytology images. In [139], a network based on the U-Net without the pooling layers combines the features from the inception network with dense blocks to make the network obtain more features with fewer parameters. The work from [140] applies a network based on the U-Net that uses a deformable convolution to improve the detection of the irregular shape of nuclei in abnormal cells. The ensemble is made using 3 different neural networks: Dense U-Net, Dense U-Net with deformable convolutions on the expansion path and a Dense U-Net with deformable convolutions on the contracting path. The authors from [141] apply an encoder and decoder architecture, like the U-Net, with a two-path parallel convo-

lutional module. The module consists of a small kernel to extract features from inside the nucleus and a large kernel to extract features from the whole cell. Also, a Feature Pyramid Network is applied to reduce the information loss on feature maps caused by the pooling layers. In [40], the authors use the U-Net model for segmentation of Argyrophilic Nucleolar Organizer Regions (AgNOR) by Silver staining technique, classifying the nuclei and the NORs as clusters and satellites. A similar method is used in [42] for nuclei classification in Papanicolaou images, where the U-Net model is applied using ResNet-34, ResNet-50, and ResNet-101 as backbones in multiple image resolutions using a binary and a multiclass approach.

### 3.7.3. Hybrid Approaches

We classified as hybrid approaches the works that use neural networks in combination with classic CV methods to segment regions of interest. The Watershed method was used along with a neural network in [142] and [143]. This method consists of pixels chosen as seeds, which then aggregate with the other pixels into *basins* according to pre-set criteria such as gray level, color, and texture.

A superpixel-based method with neural networks was applied in 3 papers. A Voronoi diagram generated with a Delaunay triangulation was used in [144]. The Delaunay triangulation divides the plane into multiple triangles, with the points as vertices, and the Voronoi algorithm uses those points to create areas where boundary lines are equidistant from neighboring points. A Multiscale Laplacian of Gaussian was used in [145], where the algorithm detects the edges of an image varying the size of the image. The paper [146] tested multiple generic segmentation methods like threshold, Hough transform, and deep learning methods as a first segmentation step, using a graph-based method to optimize the results. The approach from [147] combines a superpixel segmentation method with two Fully Convolutional Neural Networks to create an ensemble that merges the outputs from those methods. The networks used in the experiment were based on VGG-16.

## 4. Results

In this SLR, our goal was to discover the state-of-the-art of CV methods employed in computer-aided diagnosis using cytological images. We searched the Springer, IEEE, ScienceDirect, ACM, Wiley, and PubMed databases. Our research returned 431 works from which, after applying the inclusion/exclusion

criteria, remained 157 papers that we read and analyzed. During the analysis of the works, some main points of interest were mapped: the processing image method utilized, which part/organ of the body and pathology was analyzed, the stain utilized, the availability and size (considering the number of patients and images) of the used dataset, the training strategies, and the evaluation metrics applied.

Regarding the classic approaches, most of these methods are in the Space domain. The most known methods are GLCM, DWT, LBP, Watershed, and Hough Transform. There are less usual methods like Laws Texture Energy-based, Markov Random Fields, Voronoi Diagram with Delaunay Triangulation, and Multiscale Laplacian of Gaussian. For segmentation, we identified some clustering-based techniques like k-Means Clustering, DBSCAN, Fuzzy C-Means, SLIC, Superpixel Merging, Gaussian Mixture Model, and the Active Contours. There are also the morphological operations that are used, usually, in pre- and post-processing steps. In the Value domain, we could identify methods like Histogram Equalization, OTSU, and global thresholding. The less recurrent methods are the ones present in the Frequency domain, like DWT and FFT. For classification, we identified the ANN, kNN, Random Forest, Relevance Vector Machine, Naive Bayes, Decision Trees, Neighborhood Component Analysis (NCA), Random Subspace-Linear Discriminant Analysis, logistic regression model, Hierarchical Decision Model (HDA), AdaBoost, Relevance Vector Machine (RVM), Random Subspace Linear Discriminant Analysis, Neighborhood Component Analysis, and the most recurrent, the SVM.

For deep learning, most of the networks identified in this work are an application of unmodified architectures well established in the literature. For example, we could identify: VGG-11, VGG-16, VGG-19, Inception V2, Inception V3, ResNet50, ResNet101, GoogLeNet, AlexNet, DenseNet, SqueezeNet, RetinaNet and Deep Belief Networks. For object detection, we could identify YOLOv3, Faster R-CNN, MobileNet, and SSD based networks. We identified two novel architectures, the W-Net (double U-Net), and the TernausNet, composed of a modified U-Net and the DeepLabV2. For semantic segmentation, we identified the SegNet, and the U-Net, being the most used.

Figure 6 shows the works published by year and the categories into which we classified each work. Looking at the chart (Figure 6b), we can see that the number of published papers on the subject of computer-aided analysis of cytology images and the variety of methods used is growing. This fact could be explained by the growing adoption of deep learning methods (Figure 7a).

In the last few years, various network architectures for different approaches (object detection, semantic segmentation, among others) were published for the most diverse type of applications. Some of them, i.e., U-Net, were created specifically for biomedical applications.

Most of the analyzed papers use segmentation for cell localization and feature extraction (64.4%). Object detection is starting to be explored for those purposes, even though it is still not broadly employed (6.4%). Besides that, Figure 6a also shows a growing tendency of one-step methods while two-step methods, like the ones that use segmentation before classification, show a declining tendency over the last five years. Classification is used in 49.1% of the papers and it is mostly applied using single-cell images. Deep learning technologies have gained a notorious space in the computer vision community in this decade, paving the way to great advances in the machine learning area. With this in mind, it is safe to assume that in the next years the number of works using CNN technology will grow considerably. However, as shown in Figure 7b, classic approaches are still a relevant research area in the automated analysis of cytology slides, being used in 64.3% of the analyzed papers. Most of the analyzed works aim to find cell abnormalities that indicate the diagnosis of cancer. The exception are 5 papers that aim to detect pleural effusions [148, 107, 110] and vulvovaginal candidiasis [149, 150].

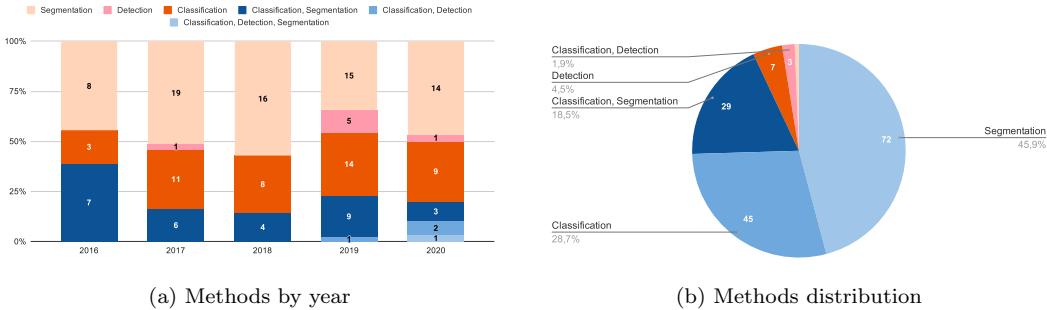


Figure 6: Number of papers by method (Segmentation, Detection, and Classification)

The most widely employed stain was Papanicolaou, present in 130 papers, followed by H&E stain (20 papers), Giemsa (5 papers), and Feulgen stain (5 papers). Pap smear datasets are used in several papers, particularly the Herlev, ISBI 2014 and 2015 challenges datasets. DCCL was the largest Papanicolaou dataset found in the analyzed works. Table 9 shows the public datasets we were able to identify. Most datasets found comprise single-cell

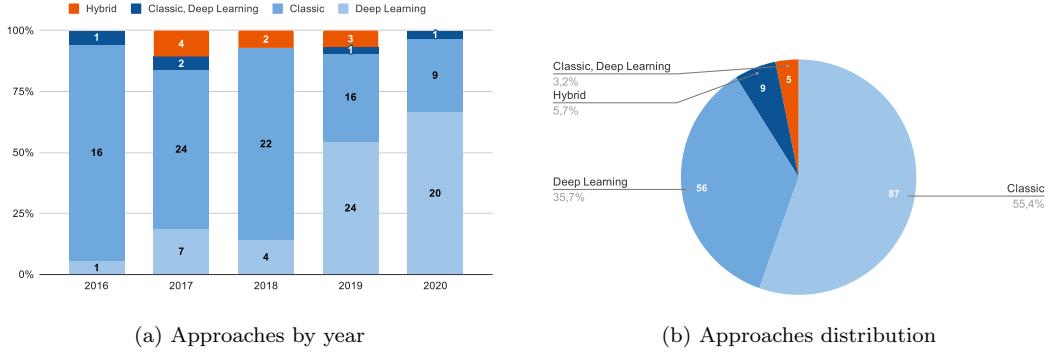


Figure 7: Number of papers by approach (Deep Learning, Classic, and Hybrid)

Pap smear cytology images and the most studied pathology is cancer (cellular proliferation).

We categorize the uses of each metric based on the type of approach the paper uses, and a supplementary analysis is available in Appendix D and Appendix E. Summarizing here an overview of the most used metrics, we identified that the most recurrent metrics used were *Accuracy*, which appears in 87 papers, and *Recall*, found in 77 papers. The third most used and the dominant metric in segmentation approaches was the *Dice Similarity Coefficient*, which was used in 55 papers, with 38 of them using segmentation approaches. The *Specificity* was the fourth most used metric, as shown in Figure 8. Even though these four metrics are the most-used, there is no common sense, since in this Systematic Literature Review (SLR) we could identify 21 different evaluation metrics.

Name	Year	Staining technique	Dataset size	Patients observed	Anatomy	Reference	Papers that have used
DTU/Herlev Pap Smear Databases	1999, 2003, 2005	Papanicolaou	917 cells	-	Cervix	[36]	[151, 146, 88, 89, 119, 102, 152, 54, 57, 153, 56, 97, 95, 65, 98, 154, 155, 133, 105, 156, 157, 140, 141, 70, 158, 86, 159, 131, 85, 118, 160, 64, 49, 120, 100, 90, 161, 59]
HEMLBC	2014	Liquid-based H&E	51 fields	21	Cervix	[32]	[146, 100]
Overlapping Cervical Cytology Image Segmentation Challenge - ISBI 2014 and 2015	2014, 2015	Liquid-based Papanicolaou	945 fields	-	Cervix	[33, 34]	[151, 162, 163, 135, 132, 164, 165, 145, 44, 166, 167, 114, 168, 169, 170, 171, 172]
-	2019	Papanicolaou	212 fields	20	Breast	[173]	[43]
-	2019	Papanicolaou	194 slides	-	Cervix	[44]	[44]
DCCL	2019	Liquid-based Papanicolaou	14,432 fields	1167	Cervix	[35]	[35]
SIPaKMeD	2018	Papanicolaou	4,049 cells	-	Cervix	[37]	[37, 86]
Perolehan Citra Kanker Serviks	2017	Papanicolaou	45 cells	-	Cervix	[38]	[174]
AGMC-TU	2018	Papanicolaou	225 cells	45	Cervix	[39]	[39]
CCAgT	2020	AgNOR	2,540 fields	3	Cervix	[40, 41]	[40]
UFSC-OCPap	2020	Papanicolaou	1,934 fields	2	Oral	[42, 175]	[42]

Table 9: Public datasets

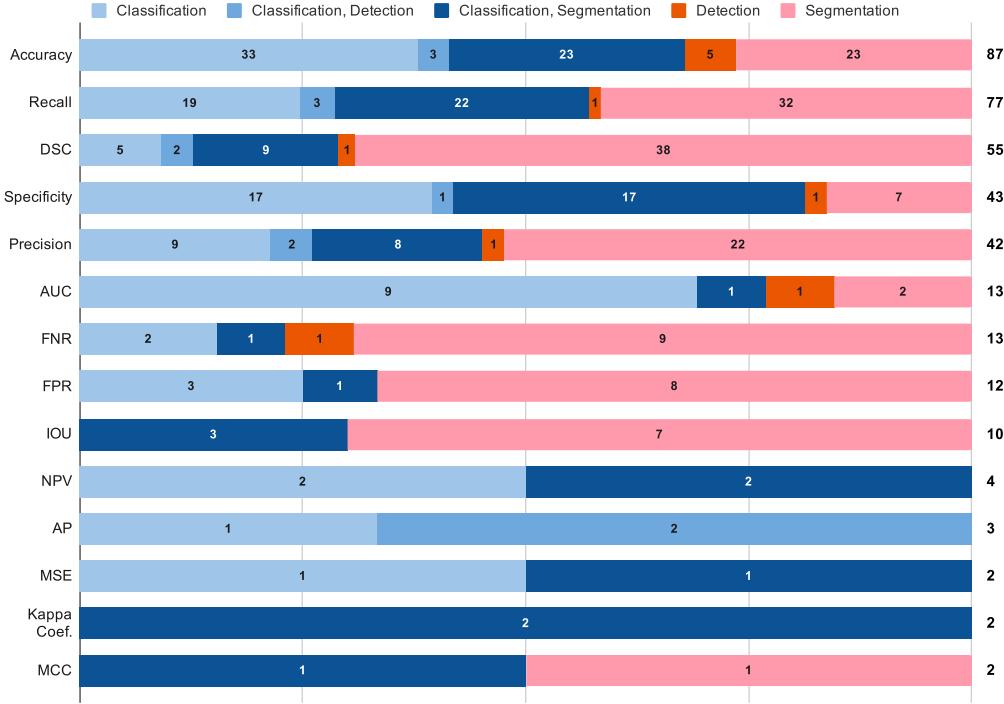


Figure 8: Evaluation metrics distribution by approach

We also identified the most commonly used training strategies, shown in Table 10 and the supplementary analysis available in Appendix C. The train-test scheme was the most used in the analyzed papers (64 papers). Of these, 41 are in articles that use ML and 23 that use classic approaches. We also point out that classics approaches tend not to use cross-validation techniques, as 25 CV papers and 17 ML papers use the full dataset for training and test.

Training Approach	Number of Papers	
	ML	Classic
Train-Test	41	23
<i>k</i> -Folds	24	7
TVT	21	0
Full Dataset	16	25

Table 10: Training and dataset handling approaches

## 5. Discussion

We observed a repetition of techniques and pipelines in groups of articles, with minor variations in the processing techniques applied. Many of these groups use similar techniques, and this led us to group them into large groups. We tried, when it made sense, to analyze the nuances and differences of each group. However, in some cases, a differentiation between papers did not make sense and that is why some works are cited only inside groups. Mostly, these groupings occurs with popular models like ANNs, SVM and Decision Trees when they are applied without great modifications.

The yearly distribution by CV approach of the works, as shown in Figure 7, indicates a crescent preponderance of DL-based methods. This is expected because the recent advances in image processing, especially in the field of modern Deep Learning techniques, have been very successful in addressing previously unsolvable problems in image detection, classification, and semantic segmentation, in different areas of application. Problems that resisted previous attempts of researchers in the field of machine learning and artificial intelligence were solved using CNN. The results achieved by CNN techniques exceed the previous ones by a large margin [14]. This way, it is natural that the researchers will try to employ these approaches to solve problems in the cytology field as well.

One of the most significant drawbacks of DL methods is the need for numerous samples to train a model. Even though we could identify some public datasets, sometimes it is necessary to gather specific data, from which there is no public dataset available, making it necessary to spend a large amount of time as a specialist to label a large amount of data to train a model. However, it is important to consider that CNNs work like a general-purpose feature extractor by adjusting the network weights using the error-backpropagation algorithm, generating complex filters that can be used to detect complex patterns [13]. Considering all the articles reviewed, about 35.7% use deep learning, as shown in Figure 6, but considering the data from the last 2 years, about 59.46% of the papers use deep learning.

Most of the networks identified in this SLR are an application of unmodified architectures that appear in the literature (VGG-16, VGG-19, Inception V3, ResNet50, ResNet101, among others). There are only two novel architectures, the W-Net (double U-Net) and the TernausNet, composed of a modified U-Net and the DeepLabV2. For semantic segmentation, the U-Net was the most recurrent architecture. Based on this information, we could

infer that the researchers are most concerned with solving the problem than create/develop new CNN architectures. In other words, this problem was not solved yet. According to this and by taking an overview look at the works found on this SLR, we could notice that most of the approaches are not ready yet to be applied on a clinical routine.

On the other side, traditional CV approaches are much more specific and highly parameter-dependent. This can explain why classical approaches found in the literature showed to be less effective when compared with DL models [176]. Classic CV techniques are strongly dependent on algorithm parameters and also on image characteristics such as contrast, noise, and resolution, making them much less robust than more state-of-the-art CNN-based approaches [177]. This suggests that CNNs are better suited for the application of computer-aided diagnosis for cytology images. Moreover, developing a hand-engineered feature extractor to use classical approaches is a difficult trial-and-error process, that involves not only mathematical knowledge about CV algorithms, but also domain knowledge that takes into account the nature and complexity of the problem. Besides that, traditional CV still seems to be an active field of research for cytology analysis applications, considering that classic and hybrid approaches are used in 64.3% of the selected papers, as shown in Figure 7.

Most of the analyzed works aim to find cell abnormalities that indicate the diagnosis of cancer. By taking into account that 152 of 157 papers presents methods to detect cancer, this seems to be the main concern of the researchers. It could be explained by the fact that some cancer types are curable if early detected. An example is the cervical cancer. It is a curable type of cancer if it could be early detected and treated properly [178]. The exception are the papers that focus on lung cells. This type of works have the goal of detect pleural effusions [148, 107, 110] and vulvovaginal candidiasis [149, 150]. Figure 7 shows the number of each type of approach found in the reviewed works. The “Classic” category stands for classical image processing methods. The category “Deep Learning” represents any deep learning-related approach, and “Hybrid” accounts all the works that used both of method types in the same pipeline. The “Classic, Deep Learning” category represents the papers that test both methods in distinct pipelines.

The most widely employed stain was Papanicolaou, present in 130 papers. However, it is necessary to point out some facts. There is a lack of standardization on its composition, coming in several formulations. The dyes that are used, their ratios, and the timing of the process could be different

on different stains [27, 179]. Some works present evidence that the analysis of cytological material using the Papanicolaou staining is not a reliable method for DNA quantification [180]. A study investigating the suitability of Papanicolaou staining and hematoxylin staining in comparison with Feulgen staining for DNA single-cell cytometry found that histograms of Feulgen stained normal squamous epithelia showed a regular DNA distribution, while Papanicolaou and hematoxylin staining showed a wide scatter of values and an increased number of values above 4c. Thus, DNA measurements from Papanicolaou-stained and hematoxylin-stained specimens are not suitable to draw diagnostic or prognostic conclusions because they cannot be interpreted reliably due to the lack of stoichiometry of nuclear staining [181].

We identified the size of the datasets and the imbalance of classes as the most recurrent problems in the analyzed papers. This indicates that there is still a lack of large and qualitatively reliable datasets comprising samples from real-world cytology examinations, as most of the approaches are based upon controlled samples (ISBI 2014/2015) or pre-processed images (Herlev). Given this, we considered DCCL as the most promising dataset and we believe it should be more explored as it was used only on its own benchmark work. It employs liquid-based cytology, as most of the datasets that we found do, which is less cost-effective than conventional cytology. We understand that this points out the necessity of good-quality datasets with images gained from conventional cytology examinations. Also, there is a lack of datasets composed of less used stains, such as the Feulgen, for example. The availability of datasets with different stains could present a direction for the researcher on defining which stain presents the best results for the attended objective.

The differences in validation strategies and metrics indicate that a simple and direct comparison between papers is not always fair because some of them, aside from distinct evaluation metrics, also use private or new datasets. One method for the measurement of the performance of object detection and semantic segmentation models that has become popular recently in the CNN literature is the Intersection Over Union (IoU) [182], as shown in Figure D.11. Cytology is an ideal field of application for the IoU because objects are all of similar size and the *large object bias* problem does not occur [183]. IoU is also considered not to be so much affected by distribution bias problems, as is the case of accuracy and similar metrics. In the cytology field, however, IoU was employed in only ten of the works we analyzed. However, it is important to emphasize the importance of using cross-validation, especially when working

with medical/cytology data, to avoid a bias in its results.

To validate our SLR strategy to avoid evidence selection bias of also searching the scientific publishers' databases directly and not only on metasearch engines such as PubMed or NCBI, we retrospectively searched on PubMed and NCBI explicitly for papers that our direct search on the scientific publishers' databases identified as relevant and, e.g., [184, 185, 186, 167, 117] are relevant publications that could not be found on neither PubMed nor NCBI. Most papers we did not find on PubMed or NCBI were ACM- or IEEE-published papers, published in CS or engineering journals, but even [117], which was published in the *Journal for Biomedical Signal Processing and Control*, an Elsevier journal available on Science Direct, could not be found on either PubMed or NCBI. We considered that these results corroborated our strategy.

## 6. Conclusions

Analyzing the papers collected in this review, we could determine that there is no publicly available dataset for many of the stains used in the described experiments. Most of the papers employ the Liquid-based Papani-colaou, even though there are variations on the stain. One impression is that many of these studies are still in a very experimental phase and not ready to be applied on a daily clinical routine. This is expected, since most of them are adopting DL methods, and this technology is relatively new, even though it advances exponentially. Looking at these works and comparing to the general tendencies that are observable in the CV community in general, we can see that, in the next few years, there will probably be a still larger adoption of DL methods for studies in the cytology image processing area, although classic Computer Vision methods are still being explored as solutions to the automation of cytology analysis.

### 6.1. Threats to Validity

This SLR was explicitly formulated in order to take into consideration only cytology papers. There exists a much vaster literature on CV for histopathology that we did not consider here. From the point of view of our knowledge on CV, we understand that many of the CV techniques used in histopathology could also be applied without much modification to cytological samples. This was a conscious delimitation of this study in order to focus on the cytology field.

## **7. Acknowledgements**

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), the Brazilian Federal Agency for Support and Evaluation of Graduate Education and by the FAPESC - Santa Catarina State Research Support Foundation. It was also supported by the Brazilian National Institute for Digital Convergence (INCoD), a research unit of the Brazilian National Institutes for Science and Technology Program (INCT) of the Brazilian National Council for Science and Technology (CNPq).

## Appendix A. Table with infos of all papers

<b>paper</b>	<b>year</b>	<b>approach</b>	<b>method</b>	<b>stain</b>	<b>dataset</b>
[51]	2016	Classification, Segmentation	Classic	Pap	
[167]	2016	Segmentation	Classic and Deep Learning	Pap	ISBI 2015
[186]	2016	Classification, Segmentation	Classic	Pap, Feulgen	
[187]	2016	Segmentation	Classic	Pap	
[56]	2016	Classification, Segmentation	Classic	Pap	Herlev
[66]	2016	Classification, Segmentation	Classic	Feulgen	
[188]	2016	Segmentation	Classic	Pap	
[64]	2016	Classification	Classic	Pap	Herlev
[189]	2016	Segmentation	Classic	Pap	
[184]	2016	Classification, Segmentation	Classic	Pap	
[115]	2016	Classification, Segmentation	Classic	H&E	
[50]	2016	Classification	Classic	Pap, H&E	
[102]	2016	Classification	Deep Learning	Pap	Herlev
[130]	2016	Segmentation	Classic	Pap, H&E	
[163]	2016	Segmentation	Classic	Pap	ISBI 2014
[185]	2016	Classification, Segmentation	Classic	Pap	
[168]	2016	Segmentation	Classic	Pap	ISBI 2014 and 2015
[190]	2016	Segmentation	Classic	Pap	
[152]	2017	Segmentation	Classic	Pap	Herlev
[191]	2017	Segmentation	Classic	Pap	
[162]	2017	Segmentation	Classic	Pap	ISBI 2014 and 2015
[171]	2017	Segmentation	Classic	Pap	ISBI 2014
[192]	2017	Segmentation	Classic	Pap	
[142]	2017	Segmentation	Hybrid	Pap	
[148]	2017	Segmentation	Classic	-	
[54]	2017	Classification, Segmentation	Classic and Deep Learning	Pap	Herlev
[193]	2017	Segmentation	Classic	Pap	
[150]	2017	Classification, Segmentation	Classic	Pap	
[114]	2017	Detection	Deep Learning	Pap	ISBI 2014
[48]	2017	Classification	Classic and Deep Learning	Pap	
[172]	2017	Segmentation	Classic	Pap	ISBI 2014
[95]	2017	Classification	Deep Learning	Pap	Herlev
[125]	2017	Segmentation	Deep Learning	Pap	
[194]	2017	Segmentation	Classic	Pap	
[52]	2017	Classification	Classic	Pap	
[100]	2017	Classification	Deep Learning	Pap, H&E	Herlev and HEMLBC
[25]	2017	Classification	Deep Learning	-	
[24]	2017	Segmentation	Deep Learning	H&E	
[146]	2017	Segmentation	Hybrid	Pap, H&E	Herlev and HEMLBC
[70]	2017	Classification	Classic	Pap	Herlev
[79]	2017	Classification	Deep Learning	H&E	
[61]	2017	Classification	Classic	Pap	
[121]	2017	Classification, Segmentation	Classic	H&E	
[28]	2017	Classification, Segmentation	Classic	H&E	
[195]	2017	Segmentation	Classic	Pap	
[49]	2017	Classification	Classic	Pap	PAP-smear
[196]	2017	Segmentation	Classic	H&E	
[105]	2017	Classification	Hybrid	Pap	Herlev
[116]	2017	Classification, Segmentation	Classic	Pap	

**Table A.11 continued from previous page**

<b>paper</b>	<b>year</b>	<b>approach</b>	<b>method</b>	<b>stain</b>	<b>dataset</b>
[144]	2017	Segmentation	Hybrid	Pap	
[197]	2017	Classification, Segmentation	Classic	Pap	
[170]	2017	Segmentation	Classic	Pap	ISBI 2014 and 2015
[174]	2017	Segmentation	Classic	Pap	Perolehan Citra Kanker Serviks
[67]	2017	Classification	Classic	Pap	
[154]	2017	Segmentation	Classic	Pap	Herlev
[118]	2018	Segmentation	Classic	Pap	Herlev
[198]	2018	Segmentation	Classic	Giemsa	
[199]	2018	Segmentation	Classic	Pap	
[58]	2018	Classification	Classic	Pap	
[200]	2018	Segmentation	Classic	Pap	
[151]	2018	Segmentation	Classic	Pap	ISBI 2014 and Herlev
[149]	2018	Classification, Segmentation	Classic	Pap	
[134]	2018	Segmentation	Deep Learning	Pap	
[74]	2018	Classification	Deep Learning	Pap	
[98]	2018	Classification	Deep Learning	Pap	Herlev
[201]	2018	Segmentation	Classic	Pap	
[202]	2018	Classification, Segmentation	Classic	Pap	
[203]	2018	Segmentation	Classic	Pap	
[57]	2018	Classification, Segmentation	Classic	Pap	Herlev
[143]	2018	Segmentation	Hybrid	H&E	
[39]	2018	Classification, Segmentation	Classic	Pap	AGMC-TU Pap-Smear
[68]	2018	Classification	Classic	Pap	
[77]	2018	Classification	Deep Learning	Giemsa	
[145]	2018	Segmentation	Hybrid	Pap	ISBI 2014 and 2015
[164]	2018	Segmentation	Classic	Pap	ISBI 2014
[204]	2018	Segmentation	Classic	Pap	
[158]	2018	Segmentation	Classic	Pap	Herlev
[157]	2018	Segmentation	Classic	Pap	Herlev
[37]	2018	Classification	Classic	Pap	SIPaKMeD
[26]	2018	Segmentation	Classic	Hematoxylin +DAB	
[63]	2018	Classification	Classic	Pap	
[53]	2018	Classification	Classic	Pap	
[120]	2018	Segmentation	Classic	Pap	Herlev
[60]	2019	Classification, Segmentation	Classic	Pap	
[111]	2019	Detection	Deep Learning	H&E	
[153]	2019	Classification, Segmentation	Classic	Pap	Herlev
[62]	2019	Classification, Segmentation	Classic	H&E	
[69]	2019	Classification	Classic	Pap	
[132]	2019	Segmentation	Deep Learning	Pap	ISBI 2014 and 2015
[205]	2019	Classification, Segmentation	Classic	-	
[83]	2019	Classification	Deep Learning	Pap	
[23]	2019	Classification, Segmentation	Classic	Feulgen	
[159]	2019	Segmentation	Classic	Pap	Herlev
[107]	2019	Detection	Deep Learning	Pap	
[113]	2019	Detection	Deep Learning	Pap	
[140]	2019	Segmentation	Deep Learning	Pap	Herlev
[161]	2019	Segmentation	Classic	Pap	Herlev
[55]	2019	Classification, Segmentation	Classic	Pap	
[86]	2019	Classification	Deep Learning	Pap	Herlev and Sipakmed
[135]	2019	Segmentation	Deep Learning	Pap	ISBI 2014
[147]	2019	Segmentation	Hybrid	Pap	
[206]	2019	Classification, Segmentation	Classic	Pap	

Table A.11 continued from previous page

paper	year	approach	method	stain	dataset
[119]	2019	Classification, Segmentation	Classic	Pap	Herlev
[155]	2019	Segmentation	Classic	Pap	Herlev
[43]	2019	Classification	Hybrid	Pap	[173]
[207]	2019	Segmentation	Classic	Pap	
[127]	2019	Segmentation	Deep Learning	H&E	
[35]	2019	Classification, Detection	Deep Learning	Pap	DCCL
[97]	2019	Classification	Deep Learning	Pap	Herlev
[89]	2019	Classification	Deep Learning	Pap	Herlev
[44]	2019	Segmentation	Deep Learning	Pap	ISBI 2014 and 2015
[109]	2019	Detection	Deep Learning	Pap	
[78]	2019	Classification	Deep Learning	H&E	
[99]	2019	Classification	Deep Learning	Pap	
[85]	2019	Classification	Deep Learning	Pap	Herlev
[137]	2019	Segmentation	Deep Learning	Pap	
[110]	2019	Detection	Deep Learning	Pap	
[104]	2019	Classification	Hybrid	Giemsa	
[65]	2019	Classification	Classic	Pap	Herlev
[101]	2019	Classification	Deep Learning	Pap	
[208]	2019	Classification	Deep Learning	Pap	
[133]	2019	Classification, Segmentation	Classic and Deep Learning	Pap	Herlev
[160]	2019	Segmentation	Classic	Pap	Herlev
[129]	2019	Segmentation	Deep Learning	Pap	
[126]	2019	Segmentation	Deep Learning	H&E	
[165]	2019	Segmentation	Classic	Pap	ISBI 2014
[84]	2019	Classification	Deep Learning	Pap	
[93]	2020	Classification	Deep Learning	Pap	
[91]	2020	Classification, Detection	Deep Learning	Pap	
[40]	2020	Segmentation	Deep Learning	AgNOR	CCAgT
[75]	2020	Classification, Detection	Deep Learning	Feulgen	
[209]	2020	Segmentation	Classic	-	
[136]	2020	Segmentation	Deep Learning	Pap	
[103]	2020	Classification	Deep Learning	Pap	
[87]	2020	Classification, Segmentation	Deep Learning	Liquid based	
[128]	2020	Segmentation	Deep Learning	Giemsa	
[131]	2020	Segmentation	Deep Learning	Pap	Herlev
[166]	2020	Segmentation	Classic	Pap	ISBI 2014
[81]	2020	Classification	Deep Learning	Pap	
[59]	2020	Classification	Classic	Pap	Herlev
[210]	2020	Segmentation	Classic	Pap	
[94]	2020	Classification	Deep Learning	Pap	
[90]	2020	Classification, Segmentation	Classic and Deep Learning	Pap	Herlev
[112]	2020	Detection	Deep Learning	Pap, H&E	
[92]	2020	Classification	Deep Learning	Pap, Feulgen	
[96]	2020	Classification	Deep Learning	Liquid based	
[141]	2020	Segmentation	Deep Learning	Pap	Herlev
[88]	2020	Classification	Deep Learning	Pap	Herlev
[139]	2020	Segmentation	Deep Learning	Pap	
[211]	2020	Segmentation	Classic	Giemsa	
[169]	2020	Segmentation	Classic	Pap	ISBI 2015
[138]	2020	Segmentation	Deep Learning	Pap, H&E	
[82]	2020	Classification	Deep Learning	Pap, H&E	
[156]	2020	Segmentation	Classic	Pap	Herlev
[42]	2020	Classification, Detection, Segmentation	Deep Learning	Pap	[175]
[117]	2020	Classification, Segmentation	Classic	H&E	
[212]	2020	Segmentation	Classic	Pap	

Table A.11: Reviewed works main points

## Appendix B. Table with details of classical segmentation papers

Paper	Objective	Pre/Pos-processing	Classification	Segmentation
[171]	Detect/Count Cells	-	-	Clustering (DB-SCAN, SLIC superpixel merging)
[169]	Detect/Count Cells	Noise reduction	-	Thresholding
[188]	Detect/Count Cells	Geometric shape constraints, iso-contours	-	Clustering (k-means)
[193]	Detect/Count Cells	Conversion to $l^*a^*b^*$ , noise reduction, morphological operations	-	Thresholding
[148]	Detect/Count Cells	Dimension reduction	-	Trainable Model (Active appearance model)
[196]	Detect/Count Cells	Conversion to grayscale, noise reduction, active contours, geometric shape constraints	-	Thresholding
[198]	Detect/Count Cells	Noise reduction, morphological operations	-	Thresholding
[151]	Detect/Count Cells	Geometric shape constraints	-	Graph-based
[26]	Detect/Count Cells	Geometric shape constraints	-	Thresholding
[207]	Detect/Count Cells	Noise reduction, convolutional filters	-	Thresholding
[117]	Detect/Count Cells, Features Extraction	Conversion to $l^*a^*b^*$ , contrast adjustments, convolutional filters, dimension reduction	SVM, Random Subspace-Linear Discriminant Analysis, Decision Tree, NCA	Clustering (k-means)
[192]	Detect/Count Cells	Convolutional filters, morphological operations	-	Thresholding
[212]	Detect/Count Cells	Conversion to grayscale, noise reduction, contrast adjustments	-	Thresholding, Clustering (k-means)
[154]	Detect/Count Cells	Noise reduction, contrast adjustments	-	Clustering
[55]	Detect/Count Cells, Segment Overlapping Cells	-	-	Clustering (Mean shift)
[56]	Detect/Count Cells	Dimension reduction	SVM	Clustering (SLIC superpixel merging)
[66]	Features Extraction	Noise reduction	Decision Tree, Logistic Regression	Thresholding
[211]	Detect/Count Cells	Conversion to grayscale, Geometric shape constraints	-	Thresholding, Graph-based
[156]	Detect/Count Cells, Features Extraction	Noise reduction, Conversion to HSV	Decision Tree, Logistic Regression, Random Forest, kNN	Clustering (k-means)
[161]	Detect/Count Cells	Conversion to grayscale, noise reduction, convolutional filters, morphological operations	-	Trainable Model (Naïve Bayes, Decision Tree, Random Forest)
[59]	Detect/Count Cells, Features Extraction, Segment Overlapping Cells	Conversion to grayscale	SVM	Thresholding
[167]	Segment Overlapping Cells	-	-	Clustering (Gaussian mixture model)
[195]	Segment Overlapping Cells	Conversion to $l^*a^*b^*$ , noise reduction, histogram equalization, morphological operations, geometric shape constraints	-	Clustering (k-means)
[162]	Segment Overlapping Cells	Noise reduction, morphological operations	-	Thresholding, Clustering (Gaussian mixture model)
[172]	Segment Overlapping Cells	Morphological operations, geometric shape constraints	-	Clustering (Fuzzy C-means)
[203]	Segment Overlapping Cells	Morphological operations	-	Thresholding

**Table B.12 continued from previous page**

Paper	Objective	Pre/Post-processing	Classification	Segmentation
[164]	Segment Overlapping Cells	Morphological operations, geometric shape constraints	-	Clustering (SLIC superpixel merging)
[165]	Segment Overlapping Cells	-	-	Clustering (SLIC superpixel merging)
[163]	Segment Overlapping Cells	Morphological operations, geometric shape constraints	-	Clustering (Fuzzy C-means)
[121]	Segment Overlapping Cells, Features Extraction	Gamma correction, morphological operations, dimension reduction	SVM, ANN, Random Forest	Thresholding, Clustering (SLIC superpixel merging)
[57]	Segment Overlapping Cells, Features Extraction	Dimension reduction, noise reduction, active contours	SVM	Graph-based
[39]	Segment Overlapping Cells, Features Extraction	Active contours	SVM, ANN, kNN, Naïve Bayes, Fuzzy C-Means, Decision Tree, Random Forest, LDA	Clustering (k-means)
[174]	Segment Overlapping Cells	Conversion to grayscale	-	Thresholding
[204]	Segment Overlapping Cells	Conversion to $l^*a^*b^*$	-	Thresholding, Clustering (k-means)
[157]	Segment Overlapping Cells	Conversion to grayscale, noise reduction, morphological operations	-	Thresholding
[168]	Segment Overlapping Cells	Conversion to grayscale, noise reduction, contrast adjustments	-	Thresholding
[187]	Features Extraction	Conversion to HSV, noise reduction	-	Thresholding
[118]	Features Extraction	Conversion to grayscale	SVM, HOG	Thresholding
[209]	Features Extraction	Conversion to grayscale, active contours	Random Forest	Thresholding
[191]	Features Extraction	Conversion to HSV, noise reduction, morphological operations	-	Thresholding
[199]	Features Extraction	Conversion to grayscale, noise reduction	-	Clustering (k-means)
[159]	Features Extraction	Conversion to grayscale, noise reduction, histogram equalization, geometric shape constraints, morphological operations, dimension reduction	Fuzzy C-Means	Thresholding
[210]	Features Extraction	Conversion to grayscale, noise reduction, morphological operations	kNN, SVM, k-means, Fuzzy C-means	Thresholding
[194]	Features Extraction	Conversion to grayscale	k-means	Thresholding
[184]	Features Extraction	Conversion to grayscale, contrast adjustments, convolutional filters, morphological operations	SVM	Thresholding
[185]	Features Extraction	Conversion to grayscale, noise reduction, morphological operations	SVM	Thresholding
[186]	Features Extraction	-	-	Graph-based
[115]	Features Extraction	Dimension reduction	SVM, ANN	Thresholding
[197]	Features Extraction	Active contours	SVM	Clustering (k-means)
[116]	Features Extraction	Noise reduction	RVM	Thresholding
[205]	Features Extraction	Conversion to $l^*a^*b^*$ , noise reduction, morphological operations, dimension reduction	SVM, Naïve Bayes	Thresholding
[23]	Features Extraction	Conversion to HSV, noise reduction, morphological operations, active contours, geometric shape constraints	Random Forest	Thresholding, Clustering (k-means)
[201]	Features Extraction	Conversion to grayscale	-	Thresholding
[28]	Features Extraction	-	Random Forest	Thresholding
[119]	Features Extraction	Conversion to grayscale, noise reduction, contrast adjustments, morphological operations, dimension reduction	SVM, kNN, Decision Tree, LDA	Clustering (k-means)

**Table B.12 continued from previous page**

Paper	Objective	Pre/Pos-processing	Classification	Segmentation
[149]	Features Extraction	Contrast adjustments	Decision Tree	Thresholding
[150]	Features Extraction	Conversion to $l^*a^*b^*$ , contrast adjustments	Decision Tree	Clustering (k-means)
[62]	Features Extraction	-	ANN	Thresholding
[51]	Features Extraction	Noise reduction, dimension reduction	SVM, ANN	Thresholding
[153]	Features Extraction	Conversion to grayscale, histogram equalization, noise reduction, geometric shape constraints, dimension reduction	Fuzzy C-means	Trainable Model
[54]	Features Extraction	Noise reduction, dimension reduction, morphological operations	SVM, ANN, Random Forest	Thresholding
[155]	Features Extraction	Conversion to grayscale, noise reduction, contrast adjustments, active contours	SVM, ANN	Thresholding
[90]	Detect/Count Cells	Morphological operations	-	Thresholding, Graph-based
[200]	Features Extraction	Conversion to grayscale, contrast adjustments, noise reduction	ANN	Thresholding
		Conversion to grayscale, contrast adjustments, histogram equalization, convolutional filters, morphological operations, dimension reduction	Random Forest	Thresholding
[158]	Features Extraction	Conversion to grayscale, convolutional filters, contrast adjustments, dimension reduction	SVM	Thresholding
[189]	Features Extraction	Conversion to $l^*a^*b^*$ , histogram equalization	SVM, kNN, LDA	Graph-based
[120]	Features Extraction	-	kNN, SVM, ANN	Thresholding
[206]	Features Extraction	Conversion to grayscale, histogram equalization, dimension reduction	Fuzzy C-means	Trainable Model
[60]	Features Extraction	Geometric shape constraints	ANN (Fuzzy)	Clustering (Fuzzy C-means)
[166]	Segment Overlapping Cells	Geometric shape constraints	-	Clustering (DB-SCAN, SLIC superpixel merging)
[152]	Detect/Count Cells	Conversion to HSV, noise reduction, geometric shape constraints	-	Clustering (Fuzzy C-means)
[190]	Detect/Count Cells	Geometric shape constraints	-	Clustering (Gaussian mixture model)
[170]	Segment Overlapping Cells	Noise reduction, contrast adjustments, histogram equalization, morphological operations, geometric shape constraints	-	Thresholding

Table B.12: Classic segmentation papers details

## Appendix C. Training and Validation Strategies

This section describes the different strategies we identified that were used to handle the datasets in order to train and test the chosen Machine Learning (ML) algorithms, and the validation strategies and methods used to validate the results.

### Appendix C.1. Train-Test Split and TTV Model

In order to avoid biased results, the dataset is commonly split into two disjunct subsets: *train* and *validation* set. The *train set* is used to train the ML algorithm and the *validation set* is used to validate the results, demonstrating the generalization capability of the trained model.

Sometimes the validation set is also used for parameter or hyperparameter tuning during training: network performance is validated at each epoch and validation results are used to optimize the training process, influencing hyperparameters such as learning rate or momentum. In this case, the validation set also has an influence on the training process, which can result in overfitting. In order to avoid that, a more robust approach is the utilization of the Train-Validation-Test (TTV) Model, where the full dataset is split into three disjunct subsets [213], as shown in Figure C.9:

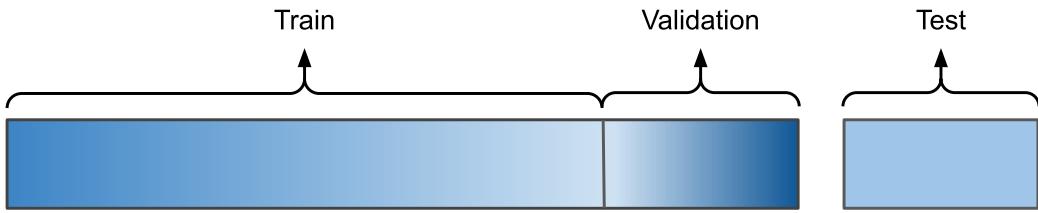


Figure C.9: A visual representation of the TTV Model

- Training set: The subset used for training the model;
- Validation set: The subset used to test the network during training and for parameter or hyperparameter tuning;
- Test set: The subset used only after training and parameter tuning to assess the performance of a fully-specified model.

The terms “validation set” and “test set” are used interchangeably in some of the papers, so we used the definitions above to perform the analysis. Also, there are 9 papers that used k-Fold for parameter tuning, instead of using distinct sets for train and validation, and these were not classified as using the TTV Model.

### Appendix C.2. *k*-Folds

*k*-folds cross-validation is a validation method where, for every  $k$  iteration, the dataset is randomly partitioned into  $k$  subsets of equal size. Then, one of these  $k$  subsets is used to validate the model while the remaining  $k - 1$  are used for training.

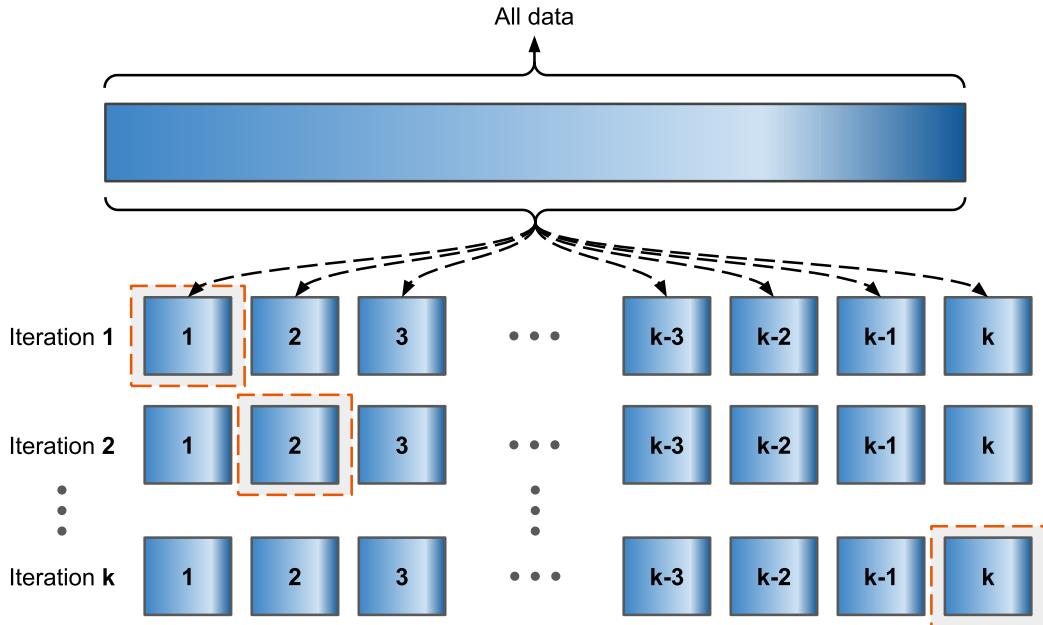


Figure C.10: A visual representation of the  $k$ -Folds cross validation

This folding process is repeated  $k$  times, resulting in each subset being used exactly one time for validation. The final result is the average of the  $k$  results obtained. The main advantage of this method is that the entire dataset is used for training and validation and each sample is used for validation exactly one time [214]. Figure C.10 shows a visual example of the  $k$ -folds cross-validation.

### Appendix D. Evaluation Metrics

We have identified several different validation strategies and metrics that were employed in order to evaluate the approaches described in the papers. Only metrics that are used in more than one of the analyzed papers are

explained. Furthermore, in the context of multi-class detection, there are many classes, which may not be uniformly distributed. This would indicate, for example, that an accuracy-based metric may introduce biases [215, 216].

The following metrics were not included due to having been used only once throughout our review: Piccard Index [88], Support [88], Relative Detection Error (RDE) [39], Aggregated Jaccard Index [39], Hausdorff distance [39], PQ [39], Spearman rank-order correlation coefficient [39] and Correlation Coefficient [39]. There is also the Mean Square Error (MSE) used at [64, 39, 42].

In order to describe each metric, it is first necessary to describe the four basic parameters used to calculate these metrics. These parameters are listed below:

- **True Positive (TP)**: also known as “hits”; number of correct positive predicted values.
- **True Negative (TN)**: also known as the “correct rejection”; is when the value is false and the prediction is equal the ground truth.
- **False Positive (FP)**: erroneous detection between of the prediction value and the ground truth.
- **False Negative (FN)**: ground truth not detected; in other words, the “misses”.

The notation applied in this work for TP, TN, FP, and FN for pixel-wise metrics is showed in the Table D.13, where:  $n_{jj}$  in TP is the number of pixels both classified and labelled as class  $j$ . Similarly,  $n_{ff}$  in TN is the number of pixels both classified and labelled as class  $f$ , the correct rejection. For false cases,  $n_{ij}$  in FP is the number of pixels that are labelled as  $i$  class, but classified as  $j$ . Finally,  $n_{ji}$  in FN is the number of pixels that are labelled as  $j$  class, but classified as  $i$  [217, 218].

These parameters can be used for classification, object detection, and semantic segmentation changing its form of implementation, and consequently changing the metrics for each desired method. For classification and object detection the implementation is similar, they work directly with the quantity of objects/items detected/classified. For semantic segmentation the implementation needs to be based on the number of pixels for each class. The equations presented in Table D.13 show both approaches, where all equations are based on these four parameters, which can also be used as a sum

of those pixels occurring, leading to a quantification of the number of pixels. This is shown visually in the Figure D.11, where we explain what are the union and intersection cases in these terms.

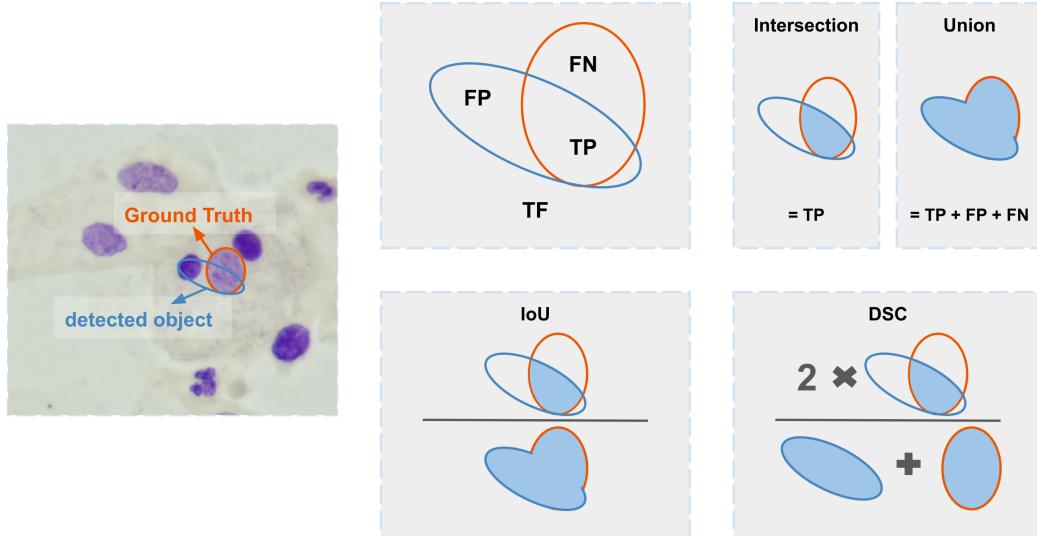


Figure D.11: Example of the application of the metrics

For better comprehension the parameters (TP, TN, FP and FN) for **object detection** are explained below:

- **True Positive (TP):** It is considered a TP when the IoU between the predicted object and the ground truth is higher than a given threshold (usually  $IoU > 0.5$ )
- **True Negative (TN):** TN is not used in object detection because it usually makes no sense to train a model with a negative case, or “background”, object
- **False Positive (FP):** When the model detects an object that is not present in the ground truth or it is present but the IoU is below a given threshold (usually 0.5).
- **False Negative (FN):** When the object is present in the ground truth and the model does not detect it.

Also, those parameters need some extra considerations: Only the prediction with the highest IoU is considered TP for one object, the others are FP; If there is an object in the ground truth, and the model predicts it but with an IoU below the threshold, it must be accounted as FP, but if there are no predictions for this object, it must be accounted for as FN; If a prediction has no corresponding ground truth, it is considered as FP.

In a similar way, the metrics can be written for both, binary- or multi-class approaches. In papers where the authors use multi-class approaches, the class-averaged metric value is usually used. However, this can be computed using also frequency-weighted values or the value for each class [215, 219].

The explanation for each metric presented in Figure 8 is enumerated below (1 to 13), and its equations are shown in the Table D.13. The *Support* metric is not listed because it represents the number of occurrences of each class in the GT.

1. **Precision:** is the ratio between the *True Positives* and the sum of *False Positives* and *True Positives* [217, 218]. This metric is also called Positive Predictive Value (PPV) and shows how accurate is the model.
2. **Recall or Sensitivity:** is the ratio between the *True Positives* and the sum of *False Negatives* and *True Positives* [217, 218], this equation is also known as of True Positive Rate (TPR). This metric shows the probability to classify the positive cases correctly.
3. **Negative Predictive Value (NPV):** is the ratio between the *True Negatives* and the sum of *False Negatives* and *True Negatives* [218]. This equation also known as True Negative Accuracy (TNA) and indicates the probability to classify negative instances correctly.
4. **False Positive Rate (FPR):** is the ratio between the *False Positives* and the sum of *False Positives* and *True Negatives* [217, 218]. In other words, it is the probability of the model to make a false positive prediction.
5. **Miss rate or False Negative Rate (FNR):** is the ratio between the *False Negative* and the sum of *False Negative* and *True Positive* [217, 218]. This is the probability of having a negative result with a positive item present in the context.
6. **Specificity:** is the ratio between the *True Negatives* and the sum of *False Positive* and *True Negatives* [217, 218]. This metric shows the probability of the model to classify the negative cases correctly.

7. **Accuracy:** is the ratio between the *True Negatives* and the sum of *False Negatives* and *True Negatives* [217]. It indicates the overall effectiveness of the model.
8. **Area under the ROC curve (AUC):** the ROC curve is plotted with TPR in the y-axis and FPR in the x-axis [218]. The ROC curve is a probability curve, so an AUC near 1 is a good measure of separability [217, 215]. This metric also measures the ability of a model to avoid false classifications.
9. **Kappa Coefficient:** is a statistic used to measure the reliability or agreement between participants for categorical items (multi-classes) [220]. In its equation the  $P_0$  is the *Accuracy*, or the total agreement probability, and  $P_c$  is the agreement probability which is due to chance [221]. This metric ranges from  $-1$  to  $+1$  and measure the inter-rater and intra-rater reliability for the classes.
10. **Matthews Correlation Coefficient (MCC):** A dataset imbalance measure that is usually used in binary classification problems. The closer to 1, the more balanced is the dataset [222].
11. **Average Precision (AP):** this metric is used mostly for object detection and is defined by the area under the Precision-Recall curve, where the precision and recall are calculated for each image. This curve is drawn with precision on the y-axis and recall on the x-axis [219]. [223, 224] describe this metric in more detail.
12. **Dice Similarity Coefficients (DSC) or F-Score:** this is a metric for semantic segmentation ans is a normalized measure of similarity, consisting of the normalized union of classes [225]. This measure is the harmonic mean of the precision and recall for a given threshold [219, 218] and is shown visually in the figure D.11.
13. **Intersection Over Union (IoU):** this is a metric for semantic segmentation and a measure based on the Jaccard Index [226] that evaluates the overlap between two regions in the image: the *ground truth region* and the *predicted region* [182, 219]. A visual representation of this metric is shown in Figure D.11. For pixel-wise approaches, this measure is the intersection of the pixel-wise classification results with the ground truth, to their union and this is the way used to measure the accuracy in model of detection objects or semantic segmentation.

Metric	Formula
TP	$\sum_{j=1}^k n_{jj}$
TN	$\sum_{f=1}^k n_{ff}$
FP	$\sum_{j=1}^k n_{ij}, \quad i \neq j$
FN	$\sum_{j=1}^k n_{ji}, \quad i \neq j$
Precision	$\frac{TP}{TP+FP} = \sum_{j=1}^k \frac{n_{jj}}{n_{ij}+n_{jj}}, \quad i \neq j$
Recall or Sensitivity	$\frac{TP}{FN+TP} = \sum_{j=1}^k \frac{n_{jj}}{n_{ji}+n_{jj}}, \quad i \neq j$
FPR	$\frac{FP}{FP+TN} = \sum_{j,f=1}^k \frac{n_{ij}}{n_{ij}+n_{ff}}, \quad i \neq j, i \neq f$
FNR	$\frac{FN}{FN+TP} = \sum_{j=1}^k \frac{n_{ji}}{n_{ji}+n_{jj}}, \quad i \neq j$
Specificity	$\frac{TN}{FP+TN} = 1 - FPR = \sum_{j,f=1}^k \frac{n_{ff}}{n_{ij}+n_{ff}}, \quad i \neq j, i \neq f$
Accuracy	$\frac{TP+TN}{FP+FN+TP+TN} = \sum_{j,f=1}^k \frac{n_{ff}+n_{jj}}{n_{jj}+n_{ff}+n_{ji}+n_{ij}}, \quad i \neq j, i \neq f$
AUC	$\frac{\text{Recall}+\text{Specificity}}{2} = \frac{1}{2} \times \left( \frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right)$
MCC	$\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$
Kappa Coefficient	$\frac{P_0 - P_c}{1 - P_c} = \frac{TP - FP - (TP+FP)(1 - 2(FP+TN))}{1 - (TP+FP)(1 - 2(FP+TN)) - FP - TN}$
DSC	$\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times TP}{FP + FN + 2TP} = \sum_{j=1}^k \frac{2 \times n_{jj}}{n_{ij} + n_{ji} + 2n_{jj}}, \quad i \neq j$
IoU	$\frac{TP}{FP + FN + TP} = \sum_{j=1}^k \frac{n_{jj}}{n_{ij} + n_{ji} + n_{jj}}, \quad i \neq j$

Table D.13: Metrics equations



## Appendix E. Table with papers that used each metric

Metric	Papers that used
Precision	[66, 102, 190, 152, 171, 142, 54, 172, 95, 25, 24, 146, 61, 195, 116, 144, 154, 57, 164, 69, 23, 140, 135, 35, 101, 133, 160, 129, 126, 165, 93, 91, 136, 87, 131, 90, 112, 96, 141, 88, 169]
Recall or Sensitivity	[51, 167, 56, 66, 188, 115, 50, 102, 130, 190, 152, 171, 142, 54, 193, 150, 172, 95, 100, 25, 24, 146, 61, 121, 195, 116, 144, 154, 199, 149, 57, 39, 145, 164, 60, 153, 62, 69, 132, 205, 83, 23, 159, 113, 140, 55, 135, 206, 119, 155, 35, 99, 85, 101, 208, 133, 160, 129, 126, 165, 93, 91, 75, 136, 87, 128, 131, 81, 94, 90, 92, 96, 141, 88, 169, 161, 59]
NPV	[61, 60, 23, 101]
FPR	[188, 130, 171, 61, 144, 199, 145, 132, 83, 159, 206, 93]
FNR	[167, 188, 130, 171, 61, 144, 199, 145, 132, 159, 206, 93, 112]
Specificity	[51, 56, 188, 115, 50, 102, 54, 150, 100, 61, 121, 116, 58, 149, 57, 39, 60, 153, 205, 83, 23, 159, 113, 55, 86, 206, 119, 155, 99, 85, 101, 208, 160, 129, 93, 75, 209, 103, 128, 81, 94, 92, 59]
Accuracy	[56, 66, 64, 189, 184, 115, 102, 130, 148, 54, 150, 114, 95, 100, 70, 79, 61, 121, 28, 49, 105, 116, 144, 197, 118, 199, 58, 200, 149, 74, 98, 57, 143, 39, 68, 77, 158, 37, 26, 63, 120, 60, 111, 153, 62, 69, 205, 83, 159, 107, 113, 55, 86, 147, 206, 119, 155, 43, 207, 35, 97, 89, 44, 109, 99, 85, 65, 160, 129, 84, 93, 91, 75, 209, 136, 103, 87, 128, 210, 94, 90, 96, 88, 82, 117, 42, 161, 59]
AUC	[162, 100, 25, 39, 68, 69, 85, 208, 160, 93, 81, 112, 92]
MCC	[147, 90]
Kappa Co-efficient	[116, 149]
Average Precision	[25, 35, 75]
DSC	[51, 167, 66, 102, 130, 168, 190, 152, 162, 171, 142, 148, 54, 172, 100, 24, 146, 121, 195, 196, 116, 144, 154, 151, 134, 39, 145, 164, 69, 132, 205, 113, 140, 86, 135, 127, 35, 137, 133, 160, 129, 126, 165, 91, 136, 128, 131, 90, 141, 88, 139, 169, 138, 40, 42, 161] 51
IoU	[121, 116, 144, 39, 147, 137, 128, 131, 156, 40, 42]

Table E.14: Metrics citations

## References

- [1] A. Diamantis, E. Magiorkinis, Pioneers of exfoliative cytology in the 19th century: the predecessors of george papanicolaou, *Cytopathology* 25 (2013) 215–224. URL: <https://doi.org/10.1111/cyt.12074>. doi:10.1111/cyt.12074.
- [2] M. Ivanovic, Overview of cytopathology procedures and techniques, in: *Cancer Treatment and Research*, Springer Berlin Heidelberg, 2013, pp. 1–12. URL: [https://doi.org/10.1007/978-3-642-38850-7\\_1](https://doi.org/10.1007/978-3-642-38850-7_1). doi:10.1007/978-3-642-38850-7\_1.
- [3] C. W. M. Bedrossian, Cytopathology: A global perspective, *Diagnostic Cytopathology* 35 (2007) 745–748. URL: <https://doi.org/10.1002/dc.20773>. doi:10.1002/dc.20773.
- [4] C. N. Chapman, C. N. Otis, From critical values to critical diagnoses, *Cancer Cytopathology* 119 (2011) 148–157. URL: <https://doi.org/10.1002/cncy.20158>. doi:10.1002/cncy.20158.
- [5] H. Li, U. Shaham, K. P. Stanton, Y. Yao, R. R. Montgomery, Y. Kluger, Gating mass cytometry data by deep learning, *Bioinformatics* 33 (2017) 3423–3430. URL: <https://doi.org/10.1093/bioinformatics/btx448>. doi:10.1093/bioinformatics/btx448.
- [6] R. Mehrotra, S. Mishra, M. Singh, M. Singh, The efficacy of oral brush biopsy with computer-assisted analysis in identifying precancerous and cancerous lesions, *Head & Neck Oncology* 3 (2011). URL: <https://doi.org/10.1186/1758-3284-3-39>. doi:10.1186/1758-3284-3-39.
- [7] M. Kandemir, F. A. Hamprecht, Computer-aided diagnosis from weak supervision: A benchmarking study, *Computerized Medical Imaging and Graphics* 42 (2015) 44–50. URL: <https://doi.org/10.1016/j.compmedimag.2014.11.010>. doi:10.1016/j.compmedimag.2014.11.010.
- [8] H. Kolles, A. V. Wangenheim, I. Niedermayer, G. H. Vince, W. Feiden, Computer assisted grading of gliomas of the astrocytoma/glioblastoma groups, *Verh Dtsch Ges Pathol* (1994) 427–431.

- [9] H. Kolles, A. V. Wangenheim, J. Rahmel, I. Niedermayer, W. Feiden, Data-driven approaches to decision making in automated tumor grading. an example of astrocytoma grading, *Anal Quant Cytol Histol* 18 (1996) 298–304.
- [10] H. Kolles, A. V. Wangenheim, G. H. Vince, W. Feiden, Automatic grading of gliomas in stereotactic biopsies. comparison of the classification results of neuronal networks and discriminant analysis, *Clinical neuropathology* 12 (1993) 253.
- [11] H. Kolles, A. V. Wangenheim, I. Niedermayer, G. H. Vince, W. Feiden, Automated grading of astrocytomas based on histomorphometric analysis of ki-67 and feulgen stained paraffin sections. classification results of neuronal networks and discriminant analysis, *Anal Cell Pathol* 8 (1995) 101–116.
- [12] H. Kolles, A. V. Wangenheim, The use of neural network technology in automated grading of astrocytoma, *Pathology - Research and Practice* 194 (1997) 254.
- [13] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [14] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (2015) 436–444. URL: <https://doi.org/10.1038/nature14539>. doi:10.1038/nature14539.
- [15] B. Kitchenham, S. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering, Technical Report, Technical report, Ver. 2.3 EBSE Technical Report. EBSE, 2007.
- [16] J. G. A. Amorim, A. Cerentini, L. A. B. Macarini, A. V. M. Matias, A. V. Wangenheim, Systematic Literature Review of Computer Vision-Aided Cytology – A Review of Classic Computer Vision and Deep Learning-Based Approaches published between January/2016 – March/2020, Technical Report 1, Federal University of Santa Catarina, 2020. DOI: 10.13140/RG.2.2.13304.67840.
- [17] A. M. Drucker, P. Fleming, A.-W. Chan, Research techniques made simple: Assessing risk of bias in systematic

- reviews, Journal of Investigative Dermatology 136 (2016) e109 – e114. URL: <http://www.sciencedirect.com/science/article/pii/S0022202X16323569>. doi:<https://doi.org/10.1016/j.jid.2016.08.021>.
- [18] W. William, A. Ware, A. H. Basaza-Ejiri, J. Obungoloch, A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images, Computer Methods and Programs in Biomedicine 164 (2018) 15–22. URL: <https://www.sciencedirect.com/science/article/pii/S0169260717307459>. doi:<https://doi.org/10.1016/j.cmpb.2018.05.034>.
  - [19] M. M. Rahaman, C. Li, X. Wu, Y. Yao, Z. Hu, T. Jiang, X. Li, S. Qi, A survey for cervical cytopathology image analysis using deep learning, IEEE Access 8 (2020) 61687–61710. doi:[10.1109/ACCESS.2020.2983186](https://doi.org/10.1109/ACCESS.2020.2983186).
  - [20] A. Sarwar, A. A. Sheikh, J. Manhas, V. Sharma, Segmentation of cervical cells for automated screening of cervical cancer: a review, Artificial Intelligence Review 53 (2020) 2341–2379. URL: <https://doi.org/10.1007/s10462-019-09735-2>. doi:[10.1007/s10462-019-09735-2](https://doi.org/10.1007/s10462-019-09735-2).
  - [21] M. Saha, R. Mukherjee, C. Chakraborty, Computer-aided diagnosis of breast cancer using cytological images: A systematic review, Tissue and Cell 48 (2016) 461–474. URL: <https://www.sciencedirect.com/science/article/pii/S0040816616300854>. doi:<https://doi.org/10.1016/j.tice.2016.07.006>.
  - [22] D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, Henry Holt and Co., Inc., USA, 1982. doi:[10.5555/1095712](https://doi.org/10.5555/1095712).
  - [23] A. H. Böcking, D. Friedrich, D. Meyer-Ebrecht, C. Zhu, A. Feider, S. Biesterfeld, Automated detection of cancer cells in effusion specimens by dna karyometry, Cancer Cytopathology 127 (2019) 18–25. URL: <https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.1002/cncy.22072>. doi:[10.1002/cncy.22072](https://doi.org/10.1002/cncy.22072).

- [24] Y. Tsukada, Y. Iwahori, K. Funahashi, M. Jose, J. Ueda, T. Iwamoto, Extraction of cell nuclei using CNN features, Procedia Computer Science 112 (2017) 1633–1640. URL: <https://doi.org/10.1016/j.procs.2017.08.255>. doi:10.1016/j.procs.2017.08.255.
- [25] H. Hu, Q. Guan, S. Chen, Z. Ji, L. Yao, Detection and recognition for life state of cell cancer using two-stage cascade CNNs, IEEE/ACM Transactions on Computational Biology and Bioinformatics (2017) 1–1. URL: <https://doi.org/10.1109/tcbb.2017.2780842>. doi:10.1109/tcbb.2017.2780842.
- [26] M. Kowal, M. Skobel, J. Korbicz, R. Monczak, Stochastic geometry for automatic assessment of ki-67 index in breast cancer preparations, in: Bioinformatics and Biomedical Engineering, Springer International Publishing, 2018, pp. 151–162. URL: [https://doi.org/10.1007/978-3-319-78759-6\\_15](https://doi.org/10.1007/978-3-319-78759-6_15). doi:10.1007/978-3-319-78759-6\_15.
- [27] G. W. Gill, Papanicolaou stain, in: Cytopreparation, Springer New York, 2012, pp. 143–189. URL: [https://doi.org/10.1007/978-1-4614-4933-1\\_10](https://doi.org/10.1007/978-1-4614-4933-1_10). doi:10.1007/978-1-4614-4933-1\_10.
- [28] S. D. Krauß, H. K. Yosef, T. Lechtenen, H. Jütte, A. Tannapfel, H. U. Käfferlein, T. Brüning, F. Roghmann, J. Noldus, S. F. El-Mashtoly, K. Gerwert, A. Mosig, Integrating spatial, morphological, and textural information for improved cell type differentiation using raman microscopy, Journal of Chemometrics 32 (2017) e2973. URL: <https://doi.org/10.1002/cem.2973>. doi:10.1002/cem.2973.
- [29] M. Derenzini, The agnors, Micron 31 (2000) 117 – 120. URL: <http://www.sciencedirect.com/science/article/pii/S0968432899000670>. doi:[https://doi.org/10.1016/S0968-4328\(99\)00067-0](https://doi.org/10.1016/S0968-4328(99)00067-0).
- [30] L. A. B. Macarini, A. von Wangenheim, F. P. Daltoé, A. S. C. Onofre, F. B. de Miranda Onofre, M. R. Stemmer, Towards a complete pipeline for segmenting nuclei in feulgen-stained images, 2020. URL: <https://arxiv.org/abs/2002.08331>. arXiv:2002.08331.
- [31] Z. Eaton-Rosen, F. Bragman, S. Ourselin, M. J. Cardoso, Improving data augmentation for medical image segmentation, in: International Conference on Medical Imaging with Deep Learning, 2018, pp. 1–3.

- [32] L. Zhang, H. Kong, C. T. Chin, S. Liu, Z. Chen, T. Wang, S. Chen, Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts, *Computerized Medical Imaging and Graphics* 38 (2014) 369–380. URL: <https://doi.org/10.1016/j.compmedimag.2014.02.001>. doi:10.1016/j.compmedimag.2014.02.001.
- [33] Z. Lu, G. Carneiro, A. P. Bradley, An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells, *IEEE Transactions on Image Processing* 24 (2015) 1261–1272. URL: <https://doi.org/10.1109/tip.2015.2389619>. doi:10.1109/tip.2015.2389619.
- [34] Z. Lu, G. Carneiro, A. P. Bradley, D. Ushizima, M. S. Nosrati, A. G. C. Bianchi, C. M. Carneiro, G. Hamarneh, Evaluation of three algorithms for the segmentation of overlapping cervical cells, *IEEE Journal of Biomedical and Health Informatics* 21 (2017) 441–450. URL: <https://doi.org/10.1109/jbhi.2016.2519686>. doi:10.1109/jbhi.2016.2519686.
- [35] C. Zhang, D. Liu, L. Wang, Y. Li, X. Chen, R. Luo, S. Che, H. Liang, Y. Li, S. Liu, D. Tu, G. Qi, P. Luo, J. Luo, DCCL: A benchmark for cervical cytology analysis, in: *Machine Learning in Medical Imaging*, Springer International Publishing, 2019, pp. 63–72. URL: [https://doi.org/10.1007/978-3-030-32692-0\\_8](https://doi.org/10.1007/978-3-030-32692-0_8). doi:10.1007/978-3-030-32692-0\_8.
- [36] J. Jantzen, G. Dounias, Analysis of pap-smear image data, in: *Proceedings of the Nature-Inspired Smart Information Systems 2nd Annual Symposium*, volume 10, NiSIS, 2006, pp. 1–11. *Nature-Inspired Smart Information Systems 2nd Annual Symposium*.
- [37] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, A. Charchanti, Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images, in: *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, 2018, pp. 3144–3148.
- [38] S. I. P. dan Klasifikasi Citra Kanker Serviks, Perolehan citra kanker serviks, 2020. URL: <http://sipk.dwiza.web.id/citra>.

- [39] M. K. Bhowmik, S. D. Roy, N. Nath, A. Datta, Nucleus region segmentation towards cervical cancer screening using AGMC-TU pap-smear dataset, in: Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence - PRAI 2018, ACM Press, 2018, pp. 44—53. URL: <https://doi.org/10.1145/3243250.3243258>. doi:10.1145/3243250.3243258.
- [40] J. G. A. Amorim, L. A. B. Macarini, A. V. Matias, A. Cerentini, F. B. D. M. Onofre, A. S. C. Onofre, A. V. Wangenheim, A novel approach on segmentation of AgNOR-stained cytology images using deep learning, in: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), IEEE, 2020, pp. 552–557. URL: <https://doi.org/10.1109/cbms49503.2020.00110>. doi:10.1109/cbms49503.2020.00110.
- [41] J. G. A. Amorim, L. A. B. Macarini, A. V. Matias, A. Cerentini, F. B. D. M. Onofre, A. S. C. Onofre, A. V. Wangenheim, Cytology dataset - CCAgT: Images of cervical cells with agnor stain technique, 2020. URL: <https://drive.google.com/drive/folders/1TBpYCv6S1ydASLauSzcsv07Wc50-WUw0?usp=sharing>.
- [42] A. V. Matias, A. Cerentini, L. A. B. Macarini, J. G. A. Amorim, F. P. Daltoé, A. von Wangenheim, Segmentation, detection and classification of cell nuclei on oral cytology samples stained with papanicolaou, in: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), 2020, pp. 53–58. doi:10.1109/CBMS49503.2020.00018.
- [43] A. R. Saikia, K. Bora, L. B. Mahanta, A. K. Das, Comparative assessment of CNN architectures for classification of breast FNAC images, *Tissue and Cell* 57 (2019) 8–14. URL: <https://doi.org/10.1016/j.tice.2019.02.001>. doi:10.1016/j.tice.2019.02.001.
- [44] F. H. Araújo, R. R. Silva, D. M. Ushizima, M. T. Rezende, C. M. Carneiro, A. G. C. Bianchi, F. N. Medeiros, Deep learning for cell image segmentation and ranking, *Computerized Medical Imaging and Graphics* 72 (2019) 13–21. URL: <https://doi.org/10.1016/j.compmedimag.2019.01.003>. doi:10.1016/j.compmedimag.2019.01.003.

- [45] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: A comprehensive review, *Neural Computation* 29 (2017) 2352–2449. URL: [https://doi.org/10.1162/neco\\_a\\_00990](https://doi.org/10.1162/neco_a_00990). doi:10.1162/neco\_a\_00990.
- [46] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, J. Schmidhuber, Flexible, high performance convolutional neural networks for image classification, in: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two, IJCAI'11, AAAI Press, 2011, p. 1237–1242. doi:10.5555/2283516.2283603.
- [47] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proceedings of the IEEE* 86 (1998) 2278–2324. URL: <https://doi.org/10.1109/5.726791>. doi:10.1109/5.726791.
- [48] C. Rasche, C. Tigăneșteanu, M. Neghină, A. Sultana, Cervical nuclei classification: Feature engineering versus deep belief network, in: Communications in Computer and Information Science, Springer International Publishing, 2017, pp. 874–885. URL: [https://doi.org/10.1007/978-3-319-60964-5\\_76](https://doi.org/10.1007/978-3-319-60964-5_76). doi:10.1007/978-3-319-60964-5\_76.
- [49] D. Lin, Z. Lin, L. Sun, K.-A. Toh, J. Cao, LLC encoded BoW features and softmax regression for microscopic image classification, in: 2017 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, 2017, pp. 1–4). URL: <https://doi.org/10.1109/iscas.2017.8050243>. doi:10.1109/iscas.2017.8050243.
- [50] S. Banerjee, S. P. K. Karri, S. Chatterjee, M. Pal, R. R. Paul, J. Chatterjee, Multimodal diagnostic segregation of oral leukoplakia and cancer, in: 2016 International Conference on Systems in Medicine and Biology (ICSMB), IEEE, 2016, pp. 67–70. URL: <https://doi.org/10.1109/icsmb.2016.7915089>. doi:10.1109/icsmb.2016.7915089.
- [51] D. Garcia-Gonzalez, M. Garcia-Silvente, E. Aguirre, A multiscale algorithm for nuclei extraction in pap smear images, *Expert Systems with Applications* 64 (2016) 512–522. URL: <https://doi.org/10.1016/j.eswa.2016.08.015>. doi:10.1016/j.eswa.2016.08.015.

- [52] L. Zhao, K. Li, J. Yin, Q. Liu, S. Wang, Complete three-phase detection framework for identifying abnormal cervical cells, *IET Image Processing* 11 (2017) 258–265. URL: <https://doi.org/10.1049/iet-ipr.2016.0788>. doi:10.1049/iet-ipr.2016.0788.
- [53] M. Arya, N. Mittal, G. Singh, Texture-based feature extraction of smear images for the detection of cervical cancer, *IET Computer Vision* 12 (2018) 1049–1059. URL: <https://doi.org/10.1049/iet-cvi.2018.5349>. doi:10.1049/iet-cvi.2018.5349.
- [54] K. Bora, M. Chowdhury, L. B. Mahanta, M. K. Kundu, A. K. Das, Automated classification of pap smear images to detect cervical dysplasia, *Computer Methods and Programs in Biomedicine* 138 (2017) 31–47. URL: <https://doi.org/10.1016/j.cmpb.2016.10.001>. doi:10.1016/j.cmpb.2016.10.001.
- [55] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, X. Hu, Automatic cell nuclei segmentation and classification of cervical pap smear images, *Biomedical Signal Processing and Control* 48 (2019) 93–103. URL: <https://doi.org/10.1016/j.bspc.2018.09.008>. doi:10.1016/j.bspc.2018.09.008.
- [56] L. Zhao, K. Li, M. Wang, J. Yin, E. Zhu, C. Wu, S. Wang, C. Zhu, Automatic cytoplasm and nuclei segmentation for color cervical smear image using an efficient gap-search MRF, *Computers in Biology and Medicine* 71 (2016) 46–56. URL: <https://doi.org/10.1016/j.combiom.2016.01.025>. doi:10.1016/j.combiom.2016.01.025.
- [57] M. A. Devi, J. Sheeba, K. S. Joseph, Neutrosophic graph cut-based segmentation scheme for efficient cervical cancer detection, *Journal of King Saud University - Computer and Information Sciences* (2018). URL: <https://doi.org/10.1016/j.jksuci.2018.09.014>. doi:10.1016/j.jksuci.2018.09.014.
- [58] T. Chankong, Automatic classifying of cervical cells using fourier spectral features, in: 2018 4th International Conference on Green Technology and Sustainable Development (GTSD), IEEE, 2018, pp. 759–762. URL: <https://ieeexplore.ieee.org/document/8595662/>. doi:10.1109/GTSD.2018.8595662.

- [59] D. Somasundaram, S. Gnanasaravanan, N. Madian, Automatic segmentation of nuclei from pap smear cell images: A step toward cervical cancer screening, International Journal of Imaging Systems and Technology (2020). URL: <https://doi.org/10.1002/ima.22444>. doi:10.1002/ima.22444.
- [60] S. Kar, D. D. Majumder, A novel approach of mathematical theory of shape and neuro-fuzzy based diagnostic analysis of cervical cancer, Pathology e Oncology Research 25 (2019) 777–790. URL: <https://doi.org/10.1007/s12253-019-00582-8>. doi:10.1007/s12253-019-00582-8.
- [61] G.-M. Makris, A. Pouliakis, C. Siristatidis, N. Margari, E. Terzakis, N. Koureas, V. Pergialiotis, N. Papantoniou, P. Karakitsos, Image analysis and multi-layer perceptron artificial neural networks for the discrimination between benign and malignant endometrial lesions, Diagnostic Cytopathology 45 (2017) 202–211. URL: <https://doi.org/10.1002/dc.23649>. doi:10.1002/dc.23649.
- [62] S. Sunny, A. Baby, B. L. James, D. Balaji, A. N. V., M. H. Rana, P. Gurpur, A. Skandarajah, M. D'Ambrosio, R. D. Ramanjinappa, S. P. Mohan, N. Raghavan, U. Kandasarma, S. N., S. Raghavan, N. Hedne, F. Koch, D. A. Fletcher, S. Selvam, M. Kollegal, P. B. N., L. Ladic, A. Suresh, H. J. Pandya, M. A. Kuriakose, A smart tele-cytology point-of-care platform for oral cancer screening, PLOS ONE 14 (2019) e0224885. URL: <https://doi.org/10.1371/journal.pone.0224885>. doi:10.1371/journal.pone.0224885.
- [63] Y. Han, S. Chen, M. Zhao, F. Shi, Suspected abnormal cervical nucleus screening based on a two-cascade classifier, in: 2018 IEEE 4th International Conference on Computer and Communications (ICCC), IEEE, 2018, pp. 2618–2624. URL: <https://doi.org/10.1109/compcomm.2018.8780633>. doi:10.1109/compcomm.2018.8780633.
- [64] M. Dumripathanachod, W. Piyawattanametha, Centralized pap test diagnosis with artificial neural network and internet of things, in: 2016 IEEE 10th International Conference on Nano/Molecular Medicine and Engineering (NANOMED), IEEE, 2016, pp. 132–135. URL: <https://doi.org/10.1109/nanomed.2016.7883577>. doi:10.1109/nanomed.2016.7883577.

- [65] S. Fekri-Ershad, Pap smear classification using combination of global significant value, texture statistical features and time series features, *Multimedia Tools and Applications* 78 (2019) 31121–31136. URL: <http://link.springer.com/10.1007/s11042-019-07937-y>. doi:10.1007/s11042-019-07937-y.
- [66] J. Su, X. Xu, Y. He, J. Song, Automatic detection of cervical cancer cells by a two-level cascade classification system, *Analytical Cellular Pathology* 2016 (2016) 1–11. URL: <https://doi.org/10.1155/2016/9535027>. doi:10.1155/2016/9535027.
- [67] S. Rodríguez-Vázquez, Support to the diagnosis of the pap test, using computer algorithms of digital image processing, in: *Advances in Computational Intelligence*, Springer International Publishing, 2017, pp. 425–436. URL: [https://doi.org/10.1007/978-3-319-62434-1\\_35](https://doi.org/10.1007/978-3-319-62434-1_35). doi:10.1007/978-3-319-62434-1\_35.
- [68] J. J. de Mesquita Sá Junior, A. R. Backes, O. M. Bruno, Pap-smear image classification using randomized neural network based signature, in: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, Springer International Publishing, 2018, pp. 677–684. URL: [https://doi.org/10.1007/978-3-319-75193-1\\_81](https://doi.org/10.1007/978-3-319-75193-1_81). doi:10.1007/978-3-319-75193-1\_81.
- [69] R. Abhinaav, D. Brindha, Abnormality detection and severity classification of cells based on features extracted from papanicolaou smear images using machine learning, in: *2019 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, 2019, pp. 1–5. URL: <https://ieeexplore.ieee.org/document/8822131/>. doi:10.1109/ICCCI.2019.8822131.
- [70] Y. Ramdhani, D. Riana, Hierarchical decision approach based on neural network and genetic algorithm method for single image classification of pap smear, in: *2017 Second International Conference on Informatics and Computing (ICIC)*, IEEE, 2017, pp. 1–6. URL: <http://ieeexplore.ieee.org/document/8280587/>. doi:10.1109/IAC.2017.8280587.
- [71] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for im-

- age recognition, 2015. URL: <https://arxiv.org/abs/1512.03385>. arXiv:1512.03385.
- [72] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014. URL: <https://arxiv.org/abs/1409.1556>. arXiv:1409.1556.
  - [73] K. Chatfield, K. Simonyan, A. Vedaldi, A. Zisserman, Return of the devil in the details: Delving deep into convolutional nets, 2014. URL: <https://arxiv.org/abs/1405.3531>. arXiv:1405.3531.
  - [74] L. D. Nguyen, D. Lin, Z. Lin, J. Cao, Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation, in: 2018 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, 2018, pp. 1–5. URL: <https://doi.org/10.1109/iscas.2018.8351550>. doi:10.1109/iscas.2018.8351550.
  - [75] Y. Xiang, W. Sun, C. Pan, M. Yan, Z. Yin, Y. Liang, A novel automation-assisted cervical cancer reading method based on convolutional neural network, *Biocybernetics and Biomedical Engineering* 40 (2020) 611–623. URL: <https://doi.org/10.1016/j.bbe.2020.01.016>. doi:10.1016/j.bbe.2020.01.016.
  - [76] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826.
  - [77] G. Dimauro, F. Girardi, M. Gelardi, V. Bevilacqua, D. Caivano, Rhino-cyt: A system for supporting the rhinologist in the analysis of nasal cytology, in: Intelligent Computing Theories and Application, Springer International Publishing, 2018, pp. 619–630. URL: [https://doi.org/10.1007/978-3-319-95933-7\\_71](https://doi.org/10.1007/978-3-319-95933-7_71). doi:10.1007/978-3-319-95933-7\_71.
  - [78] F. Bianconi, J. N. Kather, C. C. Reyes-Aldasoro, Evaluation of colour pre-processing on patch-based classification of h&e-stained images, in: Digital Pathology, Springer International Publishing, 2019, pp. 56–64. URL: [https://doi.org/10.1007/978-3-030-23937-4\\_7](https://doi.org/10.1007/978-3-030-23937-4_7). doi:10.1007/978-3-030-23937-4\_7.

- [79] H. Garud, S. P. K. Karri, D. Sheet, J. Chatterjee, M. Mahadevappa, A. K. Ray, A. Ghosh, A. K. Maity, High-magnification multi-views based classification of breast fine needle aspiration cytology cell samples using fusion of decisions from deep convolutional networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2017, pp. 76–81. URL: <https://doi.org/10.1109/cvprw.2017.115>. doi:10.1109/cvprw.2017.115.
- [80] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2015, pp. 1–9. URL: <https://doi.org/10.1109/cvpr.2015.7298594>. doi:10.1109/cvpr.2015.7298594.
- [81] D. D. E. Range, D. Dov, S. Z. Kovalsky, R. Henao, L. Carin, J. Cohen, Application of a machine learning algorithm to predict malignancy in thyroid cytopathology, *Cancer Cytopathology* 128 (2020) 287–295. URL: <https://doi.org/10.1002/cncy.22238>. doi:10.1002/cncy.22238.
- [82] M. Gupta, C. Das, A. Roy, P. Gupta, G. R. Pillai, K. Patole, Region of interest identification for cervical cancer images, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE, 2020, pp. 1293–1296. URL: <https://doi.org/10.1109/isbi45749.2020.9098587>. doi:10.1109/isbi45749.2020.9098587.
- [83] A. Teramoto, A. Yamada, Y. Kiriyama, T. Tsukamoto, K. Yan, L. Zhang, K. Imaizumi, K. Saito, H. Fujita, Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network, *Informatics in Medicine Unlocked* 16 (2019) 100205. URL: <https://doi.org/10.1016/j.imu.2019.100205>. doi:10.1016/j.imu.2019.100205.
- [84] E. Baykal, H. Dogan, M. E. Ercin, S. Ersöz, M. Ekinici, Transfer learning with pre-trained deep convolutional neural networks for serous cell classification, *Multimedia Tools and Applications* (2019). URL: <https://doi.org/10.1007/s11042-019-07821-9>. doi:10.1007/s11042-019-07821-9.

- [85] H. Lin, Y. Hu, S. Chen, J. Yao, L. Zhang, Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks, *IEEE Access* 7 (2019) 71541–71549. URL: <https://doi.org/10.1109/access.2019.2919390>. doi:10.1109/access.2019.2919390.
- [86] K. K. Gv, G. M. Reddy, Automatic classification of whole slide pap smear images using CNN with PCA based feature interpretation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2019, pp. 1074–1079. URL: <https://ieeexplore.ieee.org/document/9025606/>. doi:10.1109/CVPRW.2019.00140.
- [87] J. Deng, Y. Lu, J. Ke, An accurate neural network for cytologic whole-slide image analysis, in: Proceedings of the Australasian Computer Science Week Multiconference, ACM, 2020, pp. 1–7. URL: <https://doi.org/10.1145/3373017.3373039>. doi:10.1145/3373017.3373039.
- [88] A. Khamparia, D. Gupta, V. H. C. de Albuquerque, A. K. Sangaiah, R. H. Jhaveri, Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning, *The Journal of Supercomputing* (2020). URL: <https://doi.org/10.1007/s11227-020-03159-4>. doi:10.1007/s11227-020-03159-4.
- [89] W. Mousser, S. Ouadfel, Deep feature extraction for pap-smear image classification: A comparative study, in: Proceedings of the 2019 5th International Conference on Computer and Technology Applications, ICCTA 2019, Association for Computing Machinery, New York, NY, USA, 2019, p. 6–10. URL: <https://doi.org/10.1145/3323933.3324060>. doi:10.1145/3323933.3324060.
- [90] S. Sornapudi, G. T. Brown, Z. Xue, R. Long, L. Allen, S. Antani, Comparing Deep Learning Models for Multi-cell Classification in Liquid-based Cervical Cytology Image, *AMIA Annu Symp Proc 2019* (2019) 820–827. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7153123/>.
- [91] J. Lu, N. Sladoje, C. R. Stark, E. D. Ramqvist, J.-M. Hirsch, J. Lindblad, A deep learning based pipeline for efficient oral cancer screening on whole slide images, in: *Image Analysis and*

- Recognition, Springer International Publishing, 2020, pp. 249–261. URL: [https://doi.org/10.1007/978-3-030-50516-5\\_22](https://doi.org/10.1007/978-3-030-50516-5_22). doi:10.1007/978-3-030-50516-5\_22.
- [92] D. Gonzalez, R. L. Dietz, L. Pantanowitz, Feasibility of a deep learning algorithm to distinguish large cell neuroendocrine from small cell lung carcinoma in cytology specimens, *Cytopathology* 31 (2020) 426–431. URL: <https://doi.org/10.1111/cyt.12829>. doi:10.1111/cyt.12829.
- [93] E. Hussain, L. B. Mahanta, C. R. Das, R. K. Talukdar, A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network, *Tissue and Cell* 65 (2020) 101347. URL: <https://doi.org/10.1016/j.tice.2020.101347>. doi:10.1016/j.tice.2020.101347.
- [94] J. Martínez-Más, A. Bueno-Crespo, R. Martínez-España, M. Remezal-Solano, A. Ortiz-González, S. Ortiz-Reina, J.-P. Martínez-Cendán, Classifying papanicolaou cervical smears through a cell merger approach by deep learning technique, *Expert Systems with Applications* 160 (2020) 113707. URL: <http://www.sciencedirect.com/science/article/pii/S0957417420305315>. doi:<https://doi.org/10.1016/j.eswa.2020.113707>.
- [95] B. Taha, J. Dias, N. Werghi, Classification of cervical-cancer using pap-smear images: A convolutional neural network approach, in: *Communications in Computer and Information Science*, Springer International Publishing, 2017, pp. 261–272. URL: [https://doi.org/10.1007/978-3-319-60964-5\\_23](https://doi.org/10.1007/978-3-319-60964-5_23). doi:10.1007/978-3-319-60964-5\_23.
- [96] J. Ke, C. Liu, Y. Lu, N. Jing, X. Liang, F. Jiang, FIMIL : A high-throughput deep learning model for abnormality detection with weak annotation in microscopy images, in: *Proceedings of the Australasian Computer Science Week Multiconference*, ACM, 2020, pp. 1–6. URL: <https://doi.org/10.1145/3373017.3373051>. doi:10.1145/3373017.3373051.
- [97] S. P. B, F. Faruqi, H. K. S, R. Kudva, Deep convolution neural network for malignancy detection and classification in microscopic

- uterine cervix cell images, *Asian Pacific Journal of Cancer Prevention* 20 (2019) 3447–3456. URL: <https://doi.org/10.31557/apjcp.2019.20.11.3447>. doi:10.31557/apjcp.2019.20.11.3447.
- [98] O. U. N. Jith, K. K. Harinarayanan, S. Gautam, A. Bhavsar, A. K. Sao, DeepCerv: Deep neural network for segmentation free robust cervical cell classification, in: Computational Pathology and Ophthalmic Medical Image Analysis, Springer International Publishing, 2018, pp. 86–94. URL: [https://doi.org/10.1007/978-3-030-00949-6\\_11](https://doi.org/10.1007/978-3-030-00949-6_11). doi:10.1007/978-3-030-00949-6\_11.
- [99] K. Aljakouch, Z. Hilal, I. Daho, M. Schuler, S. D. Krauß, H. K. Yosef, J. Dierks, A. Mosig, K. Gerwert, S. F. El-Mashtoly, Fast and noninvasive diagnosis of cervical cancer by coherent anti-stokes raman scattering, *Analytical Chemistry* 91 (2019) 13900–13906. URL: <https://doi.org/10.1021/acs.analchem.9b03395>. doi:10.1021/acs.analchem.9b03395.
- [100] L. Zhang, L. Lu, I. Nogues, R. M. Summers, S. Liu, J. Yao, DeepPap: Deep convolutional networks for cervical cell classification, *IEEE Journal of Biomedical and Health Informatics* 21 (2017) 1633–1643. URL: <https://doi.org/10.1109/jbhi.2017.2705583>. doi:10.1109/jbhi.2017.2705583.
- [101] P. Sanyal, P. Ganguli, S. Barui, Performance characteristics of an artificial intelligence based on convolutional neural network for screening conventional papanicolaou-stained cervical smears, *Medical Journal Armed Forces India* (2019). URL: <https://doi.org/10.1016/j.mjafi.2019.08.001>. doi:10.1016/j.mjafi.2019.08.001.
- [102] K. Bora, M. Chowdhury, L. B. Mahanta, M. K. Kundu, A. K. Das, Pap smear image classification using convolutional neural network, in: Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, Association for Computing Machinery, New York, NY, USA, 2016, pp. 1–8. URL: <https://doi.org/10.1145/3009977.3010068>. doi:10.1145/3009977.3010068.
- [103] P. Wang, J. Wang, Y. Li, L. Li, H. Zhang, Adaptive pruning of transfer learned deep convolutional neural network for classification of cervical pap smear images, *IEEE Access* 8 (2020) 50674–50683. URL: <https://doi.org/10.1109/ACCESS.2020.2964003>.

//ieeexplore.ieee.org/document/9031411/. doi:10.1109/ACCESS.2020.2979926.

- [104] G. Dimauro, G. Ciprandi, F. Deperte, F. Girardi, E. Ladisa, S. Latrofa, M. Gelardi, Nasal cytology with deep learning techniques, International Journal of Medical Informatics 122 (2019) 13–19. URL: <https://doi.org/10.1016/j.ijmedinf.2018.11.010>. doi:10.1016/j.ijmedinf.2018.11.010.
- [105] M. Faturrahman, I. Wasito, R. Mufidah, F. D. Ghaisani, Multi feature fusion using deep belief network for automatic pap-smear cell image classification, in: 2017 International Conference on Computer, Control, Informatics and its Applications (IC3INA), IEEE, 2017, pp. 18–22. URL: <http://ieeexplore.ieee.org/document/8251733/>. doi:10.1109/IC3INA.2017.8251733.
- [106] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, 2018. arXiv:1804.02767.
- [107] B. Kilic, E. Baykal, M. Ekinci, H. Dogan, M. E. Ercin, S. Ersoz, Automated nuclei detection on pleural effusion cytopathology images using YOLOv3, in: 2019 4th International Conference on Computer Science and Engineering (UBMK), IEEE, 2019, pp. 1–5. URL: <https://doi.org/10.1109/ubmk.2019.8907125>. doi:10.1109/ubmk.2019.8907125.
- [108] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, 2017 IEEE International Conference on Computer Vision (ICCV) (2017). URL: <http://dx.doi.org/10.1109/ICCV.2017.324>. doi:10.1109/iccv.2017.324.
- [109] Du, X. Li, Q. Li, Detection and classification of cervical exfoliated cells based on faster r-CNN, in: 2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT), IEEE, 2019, pp. 52–57. URL: <https://doi.org/10.1109/icait.2019.8935931>. doi:10.1109/icait.2019.8935931.
- [110] E. Baykal, H. Dogan, M. E. Ercin, S. Ersoz, M. Ekinci, Modern convolutional object detectors for nuclei detection on pleural effusion cytology images, Multimedia Tools and Applications (2019).

URL: <https://doi.org/10.1007/s11042-019-7461-3>. doi:10.1007/s11042-019-7461-3.

- [111] S. Khan, N. Islam, Z. Jan, I. U. Din, J. J. P. C. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, *Pattern Recognition Letters* 125 (2019) 1–6. URL: <https://doi.org/10.1016/j.patrec.2019.03.022>. doi:10.1016/j.patrec.2019.03.022.
- [112] F. Su, Y. Sun, Y. Hu, P. Yuan, X. Wang, Q. Wang, J. Li, J.-F. Ji, Development and validation of a deep learning system for ascites cytopathology interpretation, *Gastric Cancer* (2020). URL: <https://doi.org/10.1007/s10120-020-01093-1>. doi:10.1007/s10120-020-01093-1.
- [113] N. Sompawong, J. Mopan, P. Pooprasert, W. Himakhun, K. Suwanarurk, J. Ngamvirojcharoen, T. Vachiramon, C. Tantibundhit, Automated pap smear cervical cancer screening using deep learning, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 7044–7048. URL: <https://doi.org/10.1109/embc.2019.8856369>. doi:10.1109/embc.2019.8856369.
- [114] R. Mufidah, I. Wasito, N. Hanifah, M. Faturrahman, F. D. Ghaisani, Automatic nucleus detection of pap smear images using stacked sparse autoencoder (ssae), in: Proceedings of the International Conference on Algorithms, Computing and Systems, ICACS '17, Association for Computing Machinery, New York, NY, USA, 2017, p. 9–13. URL: <https://doi.org/10.1145/3127942.3127946>. doi:10.1145/3127942.3127946.
- [115] L. Jeleń, A. Krzyżak, T. Fevens, M. Jeleń, Influence of feature set reduction on breast cancer malignancy classification of fine needle aspiration biopsies, *Computers in Biology and Medicine* 79 (2016) 80–91. URL: <https://doi.org/10.1016/j.combiomed.2016.10.007>. doi:10.1016/j.combiomed.2016.10.007.
- [116] A. Taneja, P. Ranjan, A. Ujlayan, Multi-cell nuclei segmentation in cervical cancer images by integrated feature vectors, *Multimedia Tools and Applications* 77 (2017) 9271–9290. URL: <https://doi.org/10.1007/s11042-017-4864-x>. doi:10.1007/s11042-017-4864-x.

- [117] A. Nawandhar, N. Kumar, V. R, L. Yamujala, Stratified squamous epithelial biopsy image classifier using machine learning and neighborhood feature selection, *Biomedical Signal Processing and Control* 55 (2020) 101671. URL: <https://doi.org/10.1016/j.bspc.2019.101671>. doi:10.1016/j.bspc.2019.101671.
- [118] S. Prum, D. O. D. Handayani, P. Boursier, Abnormal cervical cell detection using HoG descriptor and SVM classifier, in: 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), IEEE, 2018, pp. 1–6. URL: <https://ieeexplore.ieee.org/document/8776766/>. doi:10.1109/ICACCAF.2018.8776766.
- [119] K. P. Win, Y. Kitjaidure, M. P. Paing, K. Hamamoto, Cervical cancer detection and classification from pap smear images, in: Proceedings of the 2019 4th International Conference on Biomedical Imaging, Signal Processing, ICBSP '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 47–54. URL: <https://doi.org/10.1145/3366174.3366178>. doi:10.1145/3366174.3366178.
- [120] M. keymasi, V. Mishra, S. Aslan, M. M. Asem, Theoretical assessment of cervical cancer using machine learning methods based on pap-smear test, in: 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), IEEE, 2018, pp. 1367–1373. URL: <https://doi.org/10.1109/iemcon.2018.8615041>. doi:10.1109/iemcon.2018.8615041.
- [121] M. SAHA, I. ARUN, S. AGARWAL, R. AHMED, S. CHATTERJEE, C. CHAKRABORTY, Imprint cytology-based breast malignancy screening: an efficient nuclei segmentation technique, *Journal of Microscopy* 268 (2017) 155–171. URL: <https://doi.org/10.1111/jmi.12595>. doi:10.1111/jmi.12595.
- [122] Y. Guo, Y. Liu, T. Georgiou, M. S. Lew, A review of semantic segmentation using deep neural networks, *International Journal of Multimedia Information Retrieval* 7 (2017) 87–93. URL: <https://doi.org/10.1007/s13735-017-0141-z>. doi:10.1007/s13735-017-0141-z.
- [123] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, *Medical Image Com-*

- puting and Computer-Assisted Intervention – MICCAI 2015 (2015) 234–241. URL: [http://dx.doi.org/10.1007/978-3-319-24574-4\\_28](http://dx.doi.org/10.1007/978-3-319-24574-4_28). doi:10.1007/978-3-319-24574-4\_28.
- [124] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (2017) 2481–2495. URL: <http://dx.doi.org/10.1109/TPAMI.2016.2644615>. doi:10.1109/tpami.2016.2644615.
- [125] L. Zhang, M. Sonka, L. Lu, R. M. Summers, J. Yao, Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE, 2017, pp. 406–409. URL: <https://doi.org/10.1109/isbi.2017.7950548>. doi:10.1109/isbi.2017.7950548.
- [126] J. Ke, Z. Jiang, C. Liu, T. Bednarz, A. Sowmya, X. Liang, Selective detection and segmentation of cervical cells, in: Proceedings of the 2019 11th International Conference on Bioinformatics and Biomedical Technology, ACM, 2019, pp. 55–61. URL: <https://doi.org/10.1145/3340074.3340081>. doi:10.1145/3340074.3340081.
- [127] S. Das, A. Deka, Y. Iwahori, M. Bhuyan, T. Iwamoto, J. Ueda, Contour-aware residual w-net for nuclei segmentation, *Procedia Computer Science* 159 (2019) 1479–1488. URL: <https://doi.org/10.1016/j.procs.2019.09.318>. doi:10.1016/j.procs.2019.09.318.
- [128] E. B. Kablan, H. Dogan, M. E. Ercin, S. Ersoz, M. Ekinci, An ensemble of fine-tuned fully convolutional neural networks for pleural effusion cell nuclei segmentation, *Computers & Electrical Engineering* 81 (2020) 106533. URL: <https://doi.org/10.1016/j.compeleceng.2019.106533>. doi:10.1016/j.compeleceng.2019.106533.
- [129] Kurnianingsih, K. H. S. Allehaibi, L. E. Nugroho, Widayawan, L. Lazuardi, A. S. Prabuwono, T. Mantoro, Segmentation and classification of cervical cells using deep learning, *IEEE Access* 7 (2019) 116925–116941. URL: <https://doi.org/10.1109/access.2019.2936017>. doi:10.1109/access.2019.2936017.

- [130] Y. Song, J.-Z. Cheng, D. Ni, S. Chen, B. Lei, T. Wang, Segmenting overlapping cervical cell in pap smear images, in: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), IEEE, 2016, pp. 1159–1162. URL: <https://doi.org/10.1109/isbi.2016.7493472>. doi:10.1109/isbi.2016.7493472.
- [131] K. Sabeena, C. Gopakumar, R. Thampi, An improved deep convolutional model for segmentation of nucleus and cytoplasm from pap stained cell images, in: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, 2020, pp. 1274–1278. URL: <https://ieeexplore.ieee.org/document/9074244/>. doi:10.1109/ICACCS48705.2020.9074244.
- [132] T. Wan, S. Xu, C. Sang, Y. Jin, Z. Qin, Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks, Neurocomputing 365 (2019) 157–170. URL: <https://doi.org/10.1016/j.neucom.2019.06.086>. doi:10.1016/j.neucom.2019.06.086.
- [133] J. Zhao, L. Dai, M. Zhang, F. Yu, M. Li, H. Li, W. Wang, L. Zhang, PGU-net+: Progressive growing of u-net+ for automated cervical nuclei segmentation, in: Multiscale Multimodal Medical Imaging, Springer International Publishing, 2019, pp. 51–58. URL: [https://doi.org/10.1007/978-3-030-37969-8\\_7](https://doi.org/10.1007/978-3-030-37969-8_7). doi:10.1007/978-3-030-37969-8\_7.
- [134] J. Song, L. Xiao, Z. Lian, Contour-seed pairs learning-based framework for simultaneously detecting and segmenting various overlapping cells/nuclei in microscopy images, IEEE Transactions on Image Processing 27 (2018) 5759–5774. URL: <https://doi.org/10.1109/tip.2018.2857001>. doi:10.1109/tip.2018.2857001.
- [135] J. Zhang, Z. Liu, B. Du, J. He, G. Li, D. Chen, Binary tree-like network with two-path fusion attention feature for cervical cell nucleus segmentation, Computers in Biology and Medicine 108 (2019) 223–233. URL: <https://doi.org/10.1016/j.combiomed.2019.03.011>. doi:10.1016/j.combiomed.2019.03.011.
- [136] E. Hussain, L. B. Mahanta, C. R. Das, M. Choudhury, M. Chowdhury, A shape context fully convolutional neural network for segmentation

- and classification of cervical nuclei in pap smear images, *Artificial Intelligence in Medicine* 107 (2020) 101897. URL: <https://doi.org/10.1016/j.artmed.2020.101897>. doi:10.1016/j.artmed.2020.101897.
- [137] Y. Zhou, H. Chen, J. Xu, Q. Dou, P.-A. Heng, IRNet: Instance relation network for overlapping cervical cell segmentation, in: *Lecture Notes in Computer Science*, Springer International Publishing, 2019, pp. 640–648. URL: [https://doi.org/10.1007/978-3-030-32239-7\\_71](https://doi.org/10.1007/978-3-030-32239-7_71). doi:10.1007/978-3-030-32239-7\_71.
  - [138] N. A. Koohbanani, M. Jahanifar, N. Z. Tajadin, N. Rajpoot, NuClick: A deep learning framework for interactive segmentation of microscopic images, *Medical Image Analysis* 65 (2020) 101771. URL: <https://doi.org/10.1016/j.media.2020.101771>. doi:10.1016/j.media.2020.101771.
  - [139] Y. Huang, X. Li, C. Yan, L. Liu, H. Dai, MIRD-net for medical image segmentation, in: *Advances in Knowledge Discovery and Data Mining*, Springer International Publishing, 2020, pp. 207–219. URL: [https://doi.org/10.1007/978-3-030-47436-2\\_16](https://doi.org/10.1007/978-3-030-47436-2_16). doi:10.1007/978-3-030-47436-2\_16.
  - [140] J. Zhao, Q. Li, X. Li, H. Li, L. Zhang, Automated segmentation of cervical nuclei in pap smear images using deformable multi-path ensemble model, in: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2019, pp. 1514–1518. URL: <https://ieeexplore.ieee.org/document/8759262/>. doi:10.1109/ISBI.2019.8759262.
  - [141] X. Yang, J. Wu, Y. Yin, Interacting convolution with pyramid structure network for automated segmentation of cervical nuclei in pap smear images, in: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2020, pp. 499–502. URL: <https://ieeexplore.ieee.org/document/9098383/>. doi:10.1109/ISBI45749.2020.9098383.
  - [142] K. Y. Win, S. Choomchuay, K. Hamamoto, M. Raveesunthornkiat, Artificial neural network based nuclei segmentation on cytology pleural effusion images, in: *2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, IEEE, 2017, pp.

- 245–249. URL: <https://doi.org/10.1109/iciibms.2017.8279748>. doi:10.1109/iciibms.2017.8279748.
- [143] M. Kowal, M. Zejmo, J. Korbicz, Nuclei detection in cytological images using convolutional neural network and ellipse fitting algorithm, in: Artificial Intelligence and Soft Computing, Springer International Publishing, 2018, pp. 157–167. URL: [https://doi.org/10.1007/978-3-319-91262-2\\_15](https://doi.org/10.1007/978-3-319-91262-2_15). doi:10.1007/978-3-319-91262-2\_15.
- [144] A. Tareef, Y. Song, H. Huang, Y. Wang, D. Feng, M. Chen, W. Cai, Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling, Neurocomputing 248 (2017) 28–40. URL: <https://doi.org/10.1016/j.neucom.2017.01.093>. doi:10.1016/j.neucom.2017.01.093.
- [145] S. Xu, C. Sang, Y. Jin, T. Wan, Robust segmentation of overlapping cells in cervical cytology using light convolution neural network, in: Neural Information Processing, Springer International Publishing, 2018, pp. 387–397. URL: [https://doi.org/10.1007/978-3-030-04239-4\\_35](https://doi.org/10.1007/978-3-030-04239-4_35). doi:10.1007/978-3-030-04239-4\_35.
- [146] L. Zhang, H. Kong, S. Liu, T. Wang, S. Chen, M. Sonka, Graph-based segmentation of abnormal nuclei in cervical cytology, Computerized Medical Imaging and Graphics 56 (2017) 38–48. URL: <https://doi.org/10.1016/j.compmedimag.2017.01.002>. doi:10.1016/j.compmedimag.2017.01.002.
- [147] B. Harangi, J. Toth, G. Bogacsovics, D. Kupas, L. Kovacs, A. Hajdu, Cell detection on digitized pap smear images using ensemble of conventional image processing and deep learning techniques, in: 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE, 2019, pp. 38–42. URL: <https://ieeexplore.ieee.org/document/8868683/>. doi:10.1109/ISPA.2019.8868683.
- [148] E. Baykal, H. Dogan, M. Ekinci, M. E. Ercin, S. Ersoz, Automated cell nuclei segmentation in pleural effusion cytology using active appearance model, in: Computer Analysis of Images and Patterns, Springer International Publishing, 2017, pp. 59–69. URL: [https://doi.org/10.1007/978-3-319-64698-5\\_6](https://doi.org/10.1007/978-3-319-64698-5_6). doi:10.1007/978-3-319-64698-5\_6.

- [149] M. Momenzadeh, A. Vard, A. Talebi, A. Mehri Dehnavi, H. Rabbani, Computer-aided diagnosis software for vulvovaginal candidiasis detection from pap smear images, *Microscopy Research and Technique* 81 (2018) 13–21. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jemt.22951>. doi:10.1002/jemt.22951. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/jemt.22951>.
- [150] M. MOMENZADEH, M. SEHHATI, A. M. DEHNAVI, A. TALEBI, H. RABBANI, Automatic diagnosis of vulvovaginal candidiasis from pap smear images, *Journal of Microscopy* 267 (2017) 299–308. URL: <https://doi.org/10.1111/jmi.12566>. doi:10.1111/jmi.12566.
- [151] R. Saha, M. Bajger, G. Lee, Circular shape prior in efficient graph based image segmentation to segment nucleus, in: 2018 Digital Image Computing: Techniques and Applications (DICTA), IEEE, 2018, pp. 1–8. URL: <https://doi.org/10.1109/dicta.2018.8615768>. doi:10.1109/dicta.2018.8615768.
- [152] Indrabayu, A. R. Fatmasari, I. Nurtanio, A colour space based detection for cervical cancer using fuzzy c-means clustering, in: Proceedings of the 6th International Conference on Bioinformatics and Biomedical Science - ICBBS 17, ACM Press, 2017, pp. 137–141. URL: <https://doi.org/10.1145/3121138.3121196>. doi:10.1145/3121138.3121196.
- [153] W. William, A. Ware, A. H. Basaza-Ejiri, J. Obungoloch, A pap-smear analysis tool (PAT) for detection of cervical cancer from pap-smear images, *BioMedical Engineering OnLine* 18 (2019). URL: <https://doi.org/10.1186/s12938-019-0634-5>. doi:10.1186/s12938-019-0634-5.
- [154] S. Gautam, K. Gupta, A. Bhavsar, A. K. Sao, Unsupervised segmentation of cervical cell nuclei via adaptive clustering, in: Communications in Computer and Information Science, Springer International Publishing, 2017, pp. 815–826. URL: [https://doi.org/10.1007/978-3-319-60964-5\\_71](https://doi.org/10.1007/978-3-319-60964-5_71). doi:10.1007/978-3-319-60964-5\_71.
- [155] M. Arya, N. Mittal, G. Singh, Cervical cancer detection using single cell and multiple cell histopathology images, in: Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics,

- Springer Singapore, 2019, pp. 205–215. URL: [https://doi.org/10.1007/978-981-13-8300-7\\_17](https://doi.org/10.1007/978-981-13-8300-7_17). doi:10.1007/978-981-13-8300-7\_17.
- [156] H. Bandyopadhyay, M. Nasipuri, Segmentation of pap smear images for cervical cancer detection, in: 2020 IEEE Calcutta Conference (CALCON), IEEE, 2020, pp. 30–33. URL: <https://doi.org/10.1109/calcon49167.2020.9106484>. doi:10.1109/calcon49167.2020.9106484.
  - [157] S. K. Singh, R. Singh, A. Goyal, Semi-automatic segmentation of overlapping cells in pap smear image, in: 2018 4th International Conference on Computing Sciences (ICCS), IEEE, 2018, pp. 161–165. URL: <https://doi.org/10.1109/iccs.2018.00034>. doi:10.1109/iccs.2018.00034.
  - [158] V. Acharya, P. Kumar, Segmentation of pap smear images to diagnose cervical cancer types and stages, in: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2018, pp. 595–603. URL: <https://doi.org/10.1109/icacci.2018.8554911>. doi:10.1109/icacci.2018.8554911.
  - [159] W. William, A. Ware, A. H. Basaza-Ejiri, J. Obungoloch, Automated diagnosis and classification of cervical cancer from pap-smear images, in: 2019 IST-Africa Week Conference (IST-Africa), IEEE, 2019, pp. 1–11. URL: <https://ieeexplore.ieee.org/document/8764887/>. doi:10.23919/ISTAFRICA.2019.8764887.
  - [160] R. Saha, M. Bajger, G. Lee, Prior guided segmentation and nuclei feature based abnormality detection in cervical cells, in: 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, 2019, pp. 742–746. URL: <https://doi.org/10.1109/bibe.2019.00139>. doi:10.1109/bibe.2019.00139.
  - [161] W. Wasswa, J. Obungoloch, A. H. Basaza-Ejiri, A. Ware, Automated segmentation of nucleus, cytoplasm and background of cervical cells from pap-smear images using a trainable pixel level classifier, in: 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), IEEE, 2019, pp. 1–9. URL: <https://doi.org/10.1109/aipr47015.2019.9174599>. doi:10.1109/aipr47015.2019.9174599.

- [162] H. A. Phoulady, D. Goldgof, L. O. Hall, P. R. Mouton, A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images, *Computerized Medical Imaging and Graphics* 59 (2017) 38–49. URL: <https://doi.org/10.1016/j.compmedimag.2017.06.007>. doi:10.1016/j.compmedimag.2017.06.007.
- [163] R. Saha, M. Bajger, G. Lee, Spatial shape constrained fuzzy c-means (FCM) clustering for nucleus segmentation in pap smear images, in: 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Institute of Electrical and Electronics Engineers (IEEE), 2016, pp. 1–8. URL: <https://doi.org/10.1109/dicta.2016.7797086>. doi:10.1109/dicta.2016.7797086.
- [164] R. Saha, M. Bajger, G. Lee, Segmentation of cervical nuclei using SLIC and pairwise regional contrast, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 3422–3425. URL: <https://doi.org/10.1109/embc.2018.8513021>. doi:10.1109/embc.2018.8513021.
- [165] R. Saha, M. Bajger, G. Lee, SRM superpixel merging framework for precise segmentation of cervical nucleus, in: 2019 Digital Image Computing: Techniques and Applications (DICTA), IEEE, 2019, pp. 1–8. URL: <https://doi.org/10.1109/dicta47822.2019.8945887>. doi:10.1109/dicta47822.2019.8945887.
- [166] D. N. Diniz, M. J. F. Souza, C. M. Carneiro, D. M. Ushizima, F. N. S. de Medeiros, P. H. C. Oliveira, A. G. C. Bianchi, An iterated local search-based algorithm to support cell nuclei detection in pap smears test, in: Enterprise Information Systems, Springer International Publishing, 2020, pp. 78–96. URL: [https://doi.org/10.1007/978-3-030-40783-4\\_5](https://doi.org/10.1007/978-3-030-40783-4_5). doi:10.1007/978-3-030-40783-4\_5.
- [167] H. A. Phoulady, D. B. Goldgof, L. O. Hall, P. R. Mouton, A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images, in: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), IEEE, 2016, pp. 201–204. URL: <https://doi.org/10.1109/isbi.2016.7493244>. doi:10.1109/isbi.2016.7493244.

- [168] A. Bhan, G. Vyas, S. Mishra, Supervised segmentation of overlapping cervical pap smear images, in: 2016 International Conference on Signal Processing and Communication (ICSC), IEEE, 2016, pp. 225–228. URL: <https://doi.org/10.1109/icspcom.2016.7980580>. doi:10.1109/icspcom.2016.7980580.
- [169] T. Wang, J. Huang, D. Zheng, Y. He, Nucleus segmentation of cervical cytology images based on depth information, IEEE Access 8 (2020) 75846–75859. URL: <https://doi.org/10.1109/access.2020.2989369>. doi:10.1109/access.2020.2989369.
- [170] H. Nisar, L. Y. Wai, L. S. Hong, Segmentation of overlapping cells obtained from pap smear test, in: 2017 IEEE Life Sciences Conference (LSC), IEEE, 2017, pp. 254–257. URL: <https://doi.org/10.1109/lsc.2017.8268191>. doi:10.1109/lsc.2017.8268191.
- [171] P. H. C. Oliveira, G. Moreira, D. M. Ushizima, C. M. Carneiro, F. N. S. Medeiros, F. H. D. de Araujo, R. R. V. e Silva, A. G. C. Bianchi, A multi-objective approach for calibration and detection of cervical cells nuclei, in: 2017 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2017, pp. 2321–2327. URL: <https://doi.org/10.1109/cec.2017.7969586>. doi:10.1109/cec.2017.7969586.
- [172] R. Saha, M. Bajger, G. Lee, Circular shape constrained fuzzy clustering (CiscFC) for nucleus segmentation in pap smear images, Computers in Biology and Medicine 85 (2017) 13–23. URL: <https://doi.org/10.1016/j.combiomed.2017.04.008>. doi:10.1016/j.combiomed.2017.04.008.
- [173] A. R. Saikia, K. Bora, L. B. Mahanta, A. K. Das, Dataset from paper: *COMPARATIVE ASSESSMENT OF CNN ARCHITECTURES FOR CLASSIFICATION OF BREAST FNAC IMAGES*, 2019. URL: <https://onedrive.live.com/?authkey=!ABsEbvhbEc2gUFs&cid=FAD710BFDFE9935F&id=FAD710BFDFE9935F!107&parId=root&action=locate>. doi:10.1016/j.tice.2019.02.001.
- [174] D. Riana, A. N. Hidayanto, D. H. Widiantoro, T. L. R. Mengko, O. Kalsoem, Segmentation of overlapping cytoplasm and overlapped areas in pap smear images, in: 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA), IEEE,

- 2017, pp. 1–5. URL: <https://doi.org/10.1109/iisa.2017.8316386>. doi:10.1109/iisa.2017.8316386.
- [175] A. V. Matias, A. Cerentini, L. A. B. Macarini, J. G. A. Amorim, F. P. Daltoé, A. von Wangenheim, Pap smear oral sample dataset - ufsc, 2021. URL: <https://drive.google.com/drive/folders/1gT-7fKcU11q9maodENmIjOhSvTI3WBJJ?usp=sharing>.
  - [176] A. Kamilaris, F. X. Prenafeta-Boldú, Deep learning in agriculture: A survey, *Computers and Electronics in Agriculture* 147 (2018) 70–90. URL: <https://doi.org/10.1016/j.compag.2018.02.016>. doi:10.1016/j.compag.2018.02.016.
  - [177] N. O’Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, J. Walsh, Deep learning vs. traditional computer vision, in: K. Arai, S. Kapoor (Eds.), *Advances in Computer Vision*, Springer International Publishing, Cham, 2020, pp. 128–144. doi:10.1007/978-3-030-17795-9\_10.
  - [178] Y. Song, E.-L. Tan, X. Jiang, J.-Z. Cheng, D. Ni, S. Chen, B. Lei, T. Wang, Accurate cervical cell segmentation from overlapping clumps in pap smear images, *IEEE Transactions on Medical Imaging* 36 (2017) 288–300. URL: <https://doi.org/10.1109/tmi.2016.2606380>. doi:10.1109/tmi.2016.2606380.
  - [179] E. K. W. Schulte, Standardization of biological dyes and stains: pitfalls and possibilities, *Histochemistry* 95 (1991) 319–328. URL: <https://doi.org/10.1007/bf00266958>. doi:10.1007/bf00266958.
  - [180] A. M. Gurley, D. F. Hidvegi, J. W. Bacus, S. S. Bacus, Comparison of the papanicolaou and feulgen staining methods for DNA quantification by image analysis, *Cytometry* 11 (1990) 468–474. URL: <https://doi.org/10.1002/cyto.990110404>. doi:10.1002/cyto.990110404.
  - [181] S. Biesterfeld, S. Beckers, M. D. C. V. Cadenas, M. Schramm, Feulgen staining remains the gold standard for precise dna image cytometry, *Anticancer Research* 31 (2011) 53–58. URL: <https://iar.iiarjournals.org/content/31/1/53>.

- [182] Z. Zhao, P. Zheng, S. Xu, X. Wu, Object detection with deep learning: A review, CoRR abs/1807.05511 (2018). URL: <http://arxiv.org/abs/1807.05511>. arXiv:1807.05511.
- [183] G. Häger., M. Felsberg., F. Khan., Countering bias in tracking evaluations, in: Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP,, INSTICC, SciTePress, 2018, pp. 581–587. doi:10.5220/0006714805810587.
- [184] M. Arya, N. Mittal, G. Singh, Cervical cancer detection using segmentation on pap smear images, in: Proceedings of the International Conference on Informatics and Analytics - ICIA-16, ACM Press, 2016, pp. 1–5. URL: <https://doi.org/10.1145/2980258.2980311>. doi:10.1145/2980258.2980311.
- [185] R. K. Sudheesh, J. Rajan, V. S. Veena, K. Sujathan, Study of malignancy associated changes in sputum images as an indicator of lung cancer, in: 2016 IEEE Students' Technology Symposium (TechSym), IEEE, 2016, pp. 102–105. URL: <https://doi.org/10.1109/techsym.2016.7872663>. doi:10.1109/techsym.2016.7872663.
- [186] M. Toutain, A. Elmoataz, X. Desquesnes, J.-H. Pruvot, A unified geometric model for virtual slide image processing and classification, IEEE Journal of Selected Topics in Signal Processing 10 (2016) 151–160. URL: <https://doi.org/10.1109/jstsp.2015.2482941>. doi:10.1109/jstsp.2015.2482941.
- [187] F.-H. Cheng, N.-R. Hsu, Automated cell nuclei segmentation from microscopic images of cervical smear, in: 2016 International Conference on Applied System Innovation (ICASI), IEEE, 2016, pp. 1–4. URL: <https://doi.org/10.1109/icasi.2016.7539846>. doi:10.1109/icasi.2016.7539846.
- [188] M. Neghina, C. Rasche, M. Ciuc, A. Sultana, C. Tiganesteanu, Automatic detection of cervical cells in pap-smear images using polar transform and k-means segmentation, in: 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), IEEE, 2016, pp. 1–6. URL: <http://ieeexplore.ieee.org/document/7821038/>. doi:10.1109/IPTA.2016.7821038.

- [189] D. Kashyap, A. Somani, J. Shekhar, A. Bhan, M. K. Dutta, R. Burget, K. Riha, Cervical cancer detection and classification using independent level sets and multi SVMs, in: 2016 39th International Conference on Telecommunications and Signal Processing (TSP), IEEE, 2016, pp. 523–528. URL: <https://doi.org/10.1109/tsp.2016.7760935>. doi:10.1109/tsp.2016.7760935.
- [190] S. Ragothaman, S. Narasimhan, M. G. Basavaraj, R. Dewar, Unsupervised segmentation of cervical cell images using gaussian mixture model, in: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2016, pp. 70–75.
- [191] F.-H. Cheng, N.-R. Hsu, A computer-aided pap smear screening system, in: 2017 International Conference on Computational Science and Computational Intelligence (CSCI), IEEE, 2017, pp. 1807–1808. URL: <https://ieeexplore.ieee.org/document/8561085/>. doi:10.1109/CSCI.2017.320.
- [192] S. K. Singh, A. Goyal, A novel approach to segment nucleus of uterine cervix pap smear cells using watershed segmentation, in: Communications in Computer and Information Science, Springer Singapore, 2017, pp. 164–174. URL: [https://doi.org/10.1007/978-981-10-5780-9\\_15](https://doi.org/10.1007/978-981-10-5780-9_15). doi:10.1007/978-981-10-5780-9\_15.
- [193] K. Y. Win, S. Choomchuay, Automated segmentation of cell nuclei in cytology pleural fluid images using OTSU thresholding, in: 2017 International Conference on Digital Arts, Media and Technology (ICDAMT), IEEE, 2017, pp. 14–18. URL: <https://doi.org/10.1109/icdamt.2017.7904925>. doi:10.1109/icdamt.2017.7904925.
- [194] D. Riana, M. Wahyudi, A. N. Hidayanto, Comparison of nucleus and inflammatory cell detection methods on pap smear images, in: 2017 Second International Conference on Informatics and Computing (ICIC), IEEE, 2017, pp. 1–5. URL: <http://ieeexplore.ieee.org/document/8280540/>. doi:10.1109/IAC.2017.8280540.
- [195] K. Y. Win, S. Choomchuay, K. Hamamoto, K mean clustering based automated segmentation of overlapping cell nuclei in pleural effusion cytology images, in: 2017 International Conference on Advanced Technologies for Communications (ATC), IEEE, 2017, pp. 265–269.

URL: <https://doi.org/10.1109/atc.2017.8167630>. doi:10.1109/atc.2017.8167630.

- [196] H. Garud, S. P. K. Karri, D. Sheet, A. K. Maity, J. Chatterjee, M. Mahadevappa, A. K. Ray, Methods and system for segmentation of isolated nuclei in microscopic breast fine needle aspiration cytology images, in: Computer Vision, Graphics, and Image Processing, Springer International Publishing, 2017, pp. 380–392. URL: [https://doi.org/10.1007/978-3-319-68124-5\\_33](https://doi.org/10.1007/978-3-319-68124-5_33). doi:10.1007/978-3-319-68124-5\_33.
- [197] S. Dey, R. Sarkar, K. Chatterjee, P. Datta, A. Barui, S. P. Maity, Pre-cancer risk assessment in habitual smokers from DIC images of oral exfoliative cells using active contour and SVM analysis, *Tissue and Cell* 49 (2017) 296–306. URL: <https://doi.org/10.1016/j.tice.2017.01.009>. doi:10.1016/j.tice.2017.01.009.
- [198] S. U. Khan, N. Islam, Z. Jan, H. U. Shah, A. ud Din, Automated counting of cells in breast cytology images using level set method, in: 2018 IEEE 20th International Conference on High Performance Computing and Communications IEEE 16th International Conference on Smart City IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), IEEE, 2018, pp. 1578–1584. URL: <https://doi.org/10.1109/hpcc.smartcity.dss.2018.00258>. doi:10.1109/hpcc.smartcity.dss.2018.00258.
- [199] M. Sangworasil, C. Sukkasem, S. Sasivimolkul, P. Suvarnaphaet, S. Pechprasarn, R. Thongchoom, M. Janyasupab, Automated screening of cervical cancer cell images, in: 2018 11th Biomedical Engineering International Conference (BMEiCON), IEEE, 2018, pp. 1–4. URL: <https://ieeexplore.ieee.org/document/8609958/>. doi:10.1109/BMEiCON.2018.8609958.
- [200] D. Sharma, A. Bhan, A. Goyal, Cervical cancer screening in pap smear images using improved distance regularized level sets, in: 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), IEEE, 2018, pp. 1445–1448. URL: <https://doi.org/10.1109/icoei.2018.8553943>. doi:10.1109/icoei.2018.8553943.

- [201] H. Lee, M. Han, T. Yoo, C. Jung, H.-J. Son, M. Cho, Evaluation of nuclear chromatin using grayscale intensity and thresholded percentage area in liquid-based cervical cytology, *Diagnostic Cytopathology* 46 (2018) 384–389. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/dc.23906>. doi:10.1002/dc.23906.
- [202] S. Mitra, S. Dey, N. Das, S. Chakrabarty, M. Nasipuri, M. K. Naskar, Identification of benign and malignant cells from cytological images using superpixel based segmentation approach, in: *Social Transformation – Digital Way*, Springer Singapore, 2018, pp. 257–269. URL: [https://doi.org/10.1007/978-981-13-1343-1\\_24](https://doi.org/10.1007/978-981-13-1343-1_24). doi:10.1007/978-981-13-1343-1\_24.
- [203] A. Tareef, Y. Song, H. Huang, D. Feng, M. Chen, Y. Wang, W. Cai, Multi-pass fast watershed for accurate segmentation of overlapping cervical cells, *IEEE Transactions on Medical Imaging* 37 (2018) 2044–2059. URL: <https://doi.org/10.1109/tmi.2018.2815013>. doi:10.1109/tmi.2018.2815013.
- [204] D. Riana, H. Tohir, A. N. Hidayanto, Segmentation of overlapping areas on pap smear images with color features using k-means and otsu methods, in: *2018 Third International Conference on Informatics and Computing (ICIC)*, IEEE, 2018, pp. 1–5. URL: <https://doi.org/10.1109/iac.2018.8780561>. doi:10.1109/iac.2018.8780561.
- [205] S. U. Khan, N. Islam, Z. Jan, I. U. Din, A. Khan, Y. Faheem, An e-health care services framework for the detection and classification of breast cancer in breast cytology images as an IoMT application, *Future Generation Computer Systems* 98 (2019) 286–296. URL: <https://doi.org/10.1016/j.future.2019.01.033>. doi:10.1016/j.future.2019.01.033.
- [206] W. William, A. Ware, A. H. Basaza-Ejiri, J. Obungoloch, Cervical cancer classification from pap-smears using an enhanced fuzzy c-means algorithm, *Informatics in Medicine Unlocked* 14 (2019) 23 – 33. URL: <http://www.sciencedirect.com/science/article/pii/S2352914818302478>. doi:<https://doi.org/10.1016/j imu.2019.02.001>.

- [207] M. Solar, J. P. P. Gonzalez, Computational detection of cervical uterine cancer, in: 2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG), IEEE, 2019, pp. 213–217. URL: <https://doi.org/10.1109/icedeg.2019.8734400>. doi:10.1109/icedeg.2019.8734400.
- [208] A. B. Sanghvi, E. Z. Allen, K. M. Callenberg, L. Pantanowitz, Performance of an artificial intelligence algorithm for reporting urine cytopathology, *Cancer Cytopathology* 127 (2019) 658–666. URL: <https://doi.org/10.1002/cncy.22176>. doi:10.1002/cncy.22176.
- [209] D. Jia, Z. Li, C. Zhang, A parametric optimization oriented, AFSA based random forest algorithm: Application to the detection of cervical epithelial cells, *IEEE Access* 8 (2020) 64891–64905. URL: <https://ieeexplore.ieee.org/document/9051657/>. doi:10.1109/ACCESS.2020.2984657.
- [210] S. Jaya, 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), IEEE, 2020, pp. 721–726. URL: <https://ieeexplore.ieee.org/document/9156020/>.
- [211] Y. Bozhok, O. Golovko, A. Nikonenko, nPAsym: an open-source plugin for ImageJ to quantify nuclear shape asymmetry, *Computer Methods and Programs in Biomedicine* 196 (2020) 105562. URL: <https://doi.org/10.1016/j.cmpb.2020.105562>. doi:10.1016/j.cmpb.2020.105562.
- [212] M. Arya, N. Mittal, G. Singh, Three segmentation techniques to predict the dysplasia in cervical cells in the presence of debris, *Multimedia Tools and Applications* 79 (2020) 24157–24172. URL: <https://doi.org/10.1007/s11042-020-09206-9>. doi:10.1007/s11042-020-09206-9.
- [213] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 1996. URL: <https://doi.org/10.1017/cbo9780511812651>. doi:10.1017/cbo9780511812651.

- [214] G. J. McLachlan, K.-A. Do, C. Ambroise, Analyzing Microarray Gene Expression Data, John Wiley & Sons, Inc., 2004. URL: <https://doi.org/10.1002/047172842x>. doi:10.1002/047172842x.
- [215] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, *Information Processing & Management* 45 (2009) 427 – 437. URL: <http://www.sciencedirect.com/science/article/pii/S0306457309000259>. doi:<https://doi.org/10.1016/j.ipm.2009.03.002>.
- [216] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikäinen, Deep learning for generic object detection: A survey, *International Journal of Computer Vision* 128 (2020) 261 – 318. URL: <https://link.springer.com/article/10.1007/s11263-019-01247-4>. doi:<https://doi.org/10.1007/s11263-019-01247-4>.
- [217] T. Fawcett, An introduction to ROC analysis, *Pattern Recognition Letters* 27 (2006) 861–874. URL: <https://doi.org/10.1016/j.patrec.2005.10.010>. doi:10.1016/j.patrec.2005.10.010.
- [218] D. M. Powers, Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation, *Journal of Machine Learning Technologies* (2011). URL: <https://doi.org/10.9735/2229-3981>. doi:10.9735/2229-3981.
- [219] I. Ulku, E. Akagunduz, A survey on deep learning-based architectures for semantic segmentation on 2d images, *arXiv:1912.10230*, 2019. URL: <https://arxiv.org/abs/1912.10230>. arXiv:1912.10230.
- [220] M. L. McHugh, Interrater reliability: the kappa statistic, *Biochimia Medica* (2012) 276–282. URL: <https://doi.org/10.11613/bm.2012.031>. doi:10.11613/bm.2012.031.
- [221] A. Ben-David, About the relationship between ROC curves and cohens kappa, *Engineering Applications of Artificial Intelligence* 21 (2008) 874–882. URL: <https://doi.org/10.1016/j.engappai.2007.09.009>. doi:10.1016/j.engappai.2007.09.009.
- [222] B. Matthews, Comparison of the predicted and observed secondary structure of t4 phage lysozyme, *Biochimica et Biophysica Acta (BBA)*

- Protein Structure 405 (1975) 442–451. URL: [https://doi.org/10.1016/0005-2795\(75\)90109-9](https://doi.org/10.1016/0005-2795(75)90109-9). doi:10.1016/0005-2795(75)90109-9.
- [223] M. Everingham, L. V. Gool, C. K. I. Williams, J. Winn, A. Zisserman, The pascal visual object classes (VOC) challenge, International Journal of Computer Vision 88 (2009) 303–338. URL: <https://doi.org/10.1007/s11263-009-0275-4>. doi:10.1007/s11263-009-0275-4.
- [224] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, P. Dollár, Microsoft coco: Common objects in context, 2014. URL: <https://arxiv.org/abs/1405.0312>. arXiv:1405.0312.
- [225] W. Crum, O. Camara, D. Hill, Generalized overlap measures for evaluation and validation in medical image analysis, IEEE Transactions on Medical Imaging 25 (2006) 1451–1461. URL: <https://doi.org/10.1109/tmi.2006.880587>. doi:10.1109/tmi.2006.880587.
- [226] P. Jaccard, Distribution de la flore alpine dans le Bassin des Dranses et dans quelques régions voisines, volume 37, Bulletin de la Société Vaudoise des Sciences Naturellesr, Palais de Rumine, Place de la Riponne 6, 1005 Lausanne, CH, 1901.