UNIVERSIDAD NACIONAL DE COLOMBIA

Caracterización del ganglio linfático centinela para determinar la presencia de micro-metástasis

Leidy Tatiana Molano Muñoz

Universidad Nacional de Colombia
Facultad de Medicina
Bogotá, Colombia
2023

Characterization of the sentinel lymph node to determine the micro-metastases presence

Leidy Tatiana Molano Muñoz

Thesis or degree work presented as a partial requirement to apply for the title of:
Master in Biomedical Engineering

Director
Ph.D. Eduardo Romero

Research Line
Histopathology Images
Research Group
CIM@LAB - Computer Imaging and Medical Applications Laboratory

Universidad Nacional de Colombia
Faculty of Medicine
Bogotá, Colombia
2023

A mis padres Harold y Socorro,
por acomapañarme, guiarme
y apoyarme incondicionalmente en cada paso
de mi vida para poder cumplir mis objetivos
personales y académicos. Gracias por
impulsarme a perseguir mis metas y no
abandonarlas frente a las adversidades.

Acknowledgments

I sincerely express my gratitude to CIM@LAB research group, specially to Andrés Sandino, Jorge Jara, Miguel Angel Beltrán, María Jaramillo, Monica Barbosa, Jose Fuentes, Diego Bravo, Nicolás Múnera and other friends in my lab for their constant support and encouragement, critical suggestions, and their invaluable support during this process. I would like express my profound thanks to Daniel Ortega for her advices, constant support and encouragement in difficult moments those I felt weak and falter.

I would like to express my sincere thanks professor Eduardo Romero and the doctoral student Ricardo Moncayo who taught, guided and supported me during this time, Without them this thesis would not have been possible.

Abstract

The sentinel lymph node is a predictor of breast cancer aggressiveness. The hazard ratio reported was 2.14, 95% confidence interval, shows that the patient with micro-metastasis (MM) have higher probability of poorer Disease-free survival (DFS) and overall survival (OS) relative to those who are node-negative, therefore early detection of micro-metastasis analysis appears to be the approach most advantageous for patients. This work proposes an automatic detection of micro-metastasis by quantifying local cellular changes. The proposed strategy characterizes nuclei morphometry, color and texture to establish differences between MM and normal tissue. The color model is obtained from the plane $[(r-b), g]$ while texture corresponds to the Haralick's features from five different orders of the co-occurrence matrix. This description is complemented by the cellular area obtained from a conventional watershed segmentation.

Keywords: Histopathology, Detection,Breast Cancer, Micro-Metastasis.

Resumen

El ganglio linfático centinela es un predictor de la agresividad del cáncer de mama. La tasa de riesgo reportado fue de $2,14$ con un intervalo de confianza del 95 %, muestra que la paciente con micro-metástasis (MM) tiene una mayor probabilidad de peor supervivencia libre de enfermedad (SSE) y supervivencia global (SG) en relación con aquellos que tienen los ganglios negativos, por lo que la detección temprana del análisis de micro-metástasis parece ser el enforque más ventajoso para los pacientes. Este trabajo propopone una detección automática de micro-metástasis mediante la cuantificación de cambios celulares locales. La estrategia propuesta caracteriza la morfometría, el color y la textura de los núcleos para establecer diferencias entre MM y el tejido normal. El modelo de color se obtiene del plano $[(r - b), g]$ mientras que la textura corresponde a las características de Haralick de cinco órdenes diferentes de la matriz de co-ocurrencia. Esta descripción se complementa con el área celular obtenida de una segmentación de cuenca convencional.

Palabras clave: Histopatología, Detección, Cáncer de mama , Micro-Metástasis.

Contents

1 Introduction

In 2020, breast cancer disease ranked first in incidence, estimating 2.3 million of new cases. This cancer represents 11.7 % of all cancer cases, becoming the fifth cause of cancer mortality worldwide. Around 685.000 deaths are reported every year[47]. In Colombia, breast cancer has the highest cancer incidence and prevalence, representing nearly 13% of all cancers, being the leading cause of total cancer deaths [29]. Early diagnosis of this disease is crucial for effective treatment and favorable prognosis. A significantly lower probability of dying and higher survival rate is observed in patients with smaller tumors at the time of diagnosis [22]. However, breast cancer is often diagnosed in advanced stages, leading to severe outcomes. Early diagnosis and intervention are essential towards favorable outcomes [63].

1.0.1

Breast cancer is the uncontrolled growth of breast cells. During the cell life cycle, cells grow and divide as needed. When cells become abnormal or old, they usually die. Cancer starts when the cell cycle is altered, new cells are produced, and abnormal or old cells do not die when they should [5]. Breast cancer can spread from its origin to other parts of the body. Tumor cells spread from other body parts through the lymph node system (cancer cells end up in the nodes) or the bloodstream (the possibility of reaching more distant organs increases)[3]. The stage of breast cancer is crucial in selecting the best treatment for the patient. The lowest stage (Stage I and II) means that cancer does not spread. The highest stage (Stage III and IV) means cancer has spread from elsewhere [1].

The first method to analyze if cancer has spread and how far it has spread is to analyze the lymph nodes. Lymph node staging is essential for both prognosis (of early-stage disease) and treatment (for regional disease control) in patients with breast cancer[23]. This procedure consists of an axillary lymph node dissection (ALND) to perform cancer staging [24]. The sentinel lymph nodes are regional nodes that directly drain lymph from the primary tumor; this is possible because the malignant cells of the primary tumor migrate to the lymphatic vessels. Sentinel node biopsy (SNB) is the most common procedure for lymph node analysis to find metastasis (including micro-metastasis) presence when primary breast cancer is found at an early stage (I or II). Some studies show that the SNB is adequate to determine the staging of breast cancer and determine the survival rate [6, 8, 24]. This analysis has been globally adopted in the American Joint Committee on Cancer (AJCC) tumor-node-

metastasis (TNM) staging system [61], where the number the lymph nodes affected with metastasis and the size of metastasis are essential parameters.

The presence of positive axillary nodes are biomarkers to control cancer spread by evaluating local and distant regions[51]. Generally, the survival rate in patients with node-positive is up to 40% lower compared with the patients with node-negative [51, 20]. Besides, the 30% to 40% the patients with stage II and lymph nodes negative development recurrence cancer [61]. The task of finding lymph node metastasis is a statistical exercise of probability. Pathologists have succeeded in detecting metastasis in SLN, most with small-volume metastasis, such as isolated tumor cells (ITC) and micro-metastasis (MM). The micro-metastasis presence in the SLN suggests that the patient has a poorer prognosis than those who do not have the presence of micro-metastasis; that is to say, SLN is negative. The SLN diagnosis is the principal cause of determining the possible treatment [41, 44, 42].

The histopathological analysis is the gold standard for breast cancer attention and is unique to the diagnosis, the classification, and the standing of tumors [32]. However, the manual analysis in the histopathology causes intra- or inter-observers variation even for experienced pathologists [38]; for this reason, the analysis results can be inaccurate evaluation[14]. Hematoxylin, Eosin (H&E), and immunohistochemical (IHC) stained tissue are available histology images for detecting and quantifying metastasis. Nevertheless, these analyses are expensive, time-consuming tasks and highly a pathologist's expertise. Besides, during the examination, some micro-metastasis are unnoticed, becoming complex.

The digital microscopy is an useful alternative to conventional light microscopy [43]. The digital tissue section, called Whole Slide Image (WSI), opens the door to applying different technics of analysis to the application of artificial intelligence (AI) and machine learning techniques[32]. This image modality has allowed to develop some methods for automatic whole slide image (WSI) analysis that can be perform to detect SNL metastasis. In this work, we perform an automatic strategy to identify micro-metastasis, in Hemathoxylin and Eosin (H&E) images, that allow to quantify the disease and potentially support a personalized treatment.

1.1 Breast Cancer disease

The breast is the organ located over the upper ribs and chest muscles. It comprises three main parts: lobules, ducts, and connective tissue,as show in the **1-1**. The lobules, which are a gland, produce milk for the baby. The ducts are small canals that come from the lobules and carry the milk to the nipple and. The connective tissue (which consists of fibrous and fatty tissue) surrounds and holds everything together.

Cancer is produced by changes in the cells and mutations in the genes that regulate the growth of cells. Breast cancer is when mammary cells acquire the ability to divide without

any control or order, thus producing more of the same cells and generating a tumor [21]. The tumor can be benign (is not dangerous to health) or malignant (potentially dangerous). The begging tumor is not considered cancerous because the cells have a normal appearance, grow slowly, and do not invade nearby tissues or spread to other body parts. The malign tumor is cancerous and more likely to disseminate to other body parts [5]. The principal breast cancer detection methods are self-examination, clinical examination, and mammography. The possible treatment depends on the biopsy and the tumor pathology. However, the treatment can be individualized because the biology of each tumor is different [2].

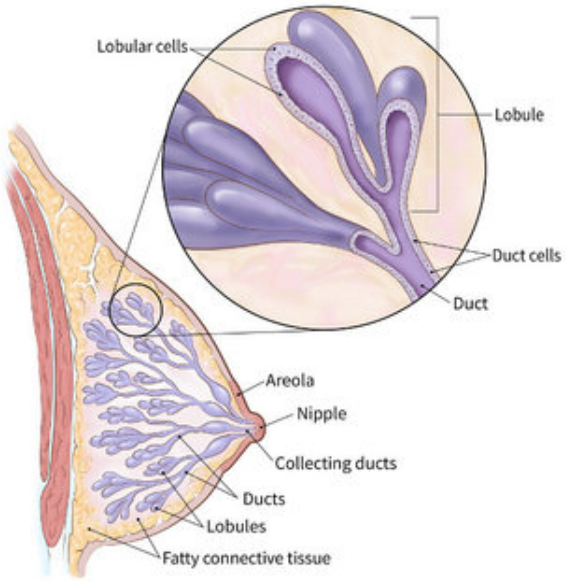


Figure 1-1: Breast anatomy

taken from: www.cancer.org

Breast cancer can start in different parts of the breast, and the type of breast cancer depends on the kind of cell affected. There are many types of breast cancer and many different ways to describe them. The most common cancers are ductal (cancer cells originate in the ductal) and lobular carcinoma (cancer cells originate in the lobular). However, the type of breast cancer also depends on whether the cancer cell invades other tissues. The term *invasive* or $infiltrating$ defines any breast cancer that has spread (invaded) into the surrounding breast tissue, and the term $in-situ$ refers to the fact that cancer has not grown into the rest of the breast tissue [4].

1.1.1 Breast Cancer Stages

The prognosis of breast cancer usually expresses the survivor rate. This survivor rate is based on the cancer stage at the diagnosis. The stage of breast cancer is essential to determine the

best treatment and describe how much cancer the body has [1]. The stage is determined by features such as size and hormonal receptors. However, the principal analysis to determine this is using the TNM classifier adopted by the American Joint Committee on Cancer (AJCC) [61].
The medical may refer to the stage of the breast as stages I and II as early stages and the stage IIB and III as the advance stage.

- **Stage 0:** This stage describes breast cancer as not invasive and is restricted to the ducts or lobules. There is no evidence of cancer cells outside the area where they originated.
- **Stage I:** The cancer cells invading the normal mammary tissue. The stage I is divide in two categories as IA (the tumor is invasive but has not spread to the lymph nodes) and IB (The tumor is propagated to the lymph node, and tumor size is between 0.2mm and 2 mm).
- **Stage IIA:** This state meets conditions such as:
  - There is no tumor in the breast, but the cancer is disseminated from the lymph node (1 at 3 lymph nodes).
  - The tumor size is smaller 20mm, and it is disseminated from the lymph nodes.
  - The tumor size is between 20mm and 50mm, and it is not disseminated from the lymph nodes.
- **Stage IIB:** This state meets conditions such as:
  - The tumor size is between 20mm and 50mm,and the cancer is disseminated from the lymph node (1 at 3 lymph nodes).
  - The tumor size is larger 50mm, and it is not disseminated from the lymph nodes.
- **Stage IIIA:** The cancer is disseminated from the lymph nodes (4 at 9 lymph nodes) independently of the tumor size. Alternatively, the tumor is larger by 50 mm but only affects 1 at 3 lymph nodes.
- **Stage IIIB:** The tumor is disseminated from the thorax wall or skin of the breast and produces inflammation or an ulcer.
- **Stage IIIC:** The cancer is disseminated from the lymph node in more that 10 lymph nodes.
- **Stage IV (Metastastic):** The tumor can be any size and disseminated from other organs, such as the lung, bone, skin, brain, etc.
- **Recurrence:** It is cancer that appears after treatment.

1.1.2 TNM Staging System

Breast cancer can spread outside the breast through blood and lymph vessels. When breast cancer spreads to other body parts, it is said to have metastasized [5]. The most common medical tool to describe the stage of the breast cancer is the TNM system: $Tumor(T)$ describes the primary tumor, $Node(N)$ indicates if the tumor propagates to the lymph nodes, and $Metastasis(M)$ indicator if the cancer was metastasis to other body parts [1].

- **Primary tumor (T):** To determine how much and where the cancer is in the body, the medical analyse the primary tumor. Analyze the size and the localization of the tumor.
  - **TX:** This means that there is no information about the primary tumor or that it cannot be measured.
  - **T0:** This means that there is no evidence of cancer in the mammary gland.
  - **Tis** This means that the cancer cells are only growing in the layer of cells where they started, without growing into deeper layers. This may also be called in-situ cancer.
  - A number after the T (such as T1, T2, T3, or T4) might describe the tumor size and amount of spread into nearby structures. The higher the T number, the larger the tumor and/or the more it has grown into nearby tissues.
- **Lymph Nodes (N):** This category describes the lymph node close to the primary tumor, with the proposal to determine if cancer disseminates to them before arriving at other body parts.
  - **NX:** This means it does not have information about the lymph nodes close or that cannot be evaluated.
  - **N0:** This means that the lymph nodes do not have cancer.
  - A number after the N (such as N1, N2, N3, or N4) might describe how many lymph nodes close are affected. The higher the N number, the more significant the propagation of cancer in the lymph nodes.
- **Metastasis (M):** This category examines other body parts whit the proposal to determine if cancer disseminates. The primary cancer propagation from the other body parts is known as *metastasis*.
  - **M0:** This means that primary cancer has not spread.
  - **M1:** This means that primary cancer has spread from the other body part.

1.2 Importance of Sentinel lymph node in Breast Cancer

The lymphatic system is a vital part of the immune system; this is similar to the circulatory system and a network of lymph nodes connected by lymphatic vessels that transport lymph throughout the body. The lymph is formed from the fluid that seeps through the thin walls of the capillaries into the body's tissues; this fluid contains oxygen, proteins, and other nutrients that nourish the tissues. Lymph also carries foreign substances (such as bacteria), cancer cells, and dead or damaged cells that may be present in the tissues of the lymphatic vessels and organs for removal. All substances carried by lymph pass through at least one lymph node, where foreign substances are filtered out and destroyed before the fluid returns to the bloodstream. Therefore, the main functions of the lymphatic system are to remove damaged cells from the body and to provide protection against the spread of infections and cancer [31].

Breast cancer usually spreads to the axillary lymph nodes (ALN) before any other location. Given their proximity to the breasts, show the **1-2**, these lymph nodes are a common site for spreading breast cancer. If cancer cells have spread to the lymph nodes, there is a greater chance that the cells have traveled through the lymphatic system and have spread (metastasized) to other parts of the body. Axillary lymph node dissection (ALND) is an essential component of the surgical treatment of breast cancer, as it is a cancer staging procedure. However, this procedure is associated with unpleasant side effects; sentinel lymph node biopsy (SLNB) was implemented to reduce side effects; this analysis reduces the adverse effect with the proposed staging of breast cancer [50]. Historically, node-negative patients underwent SLNB, while node-positive patients underwent ALND. The sentinel lymph node is the first lymph node where the cancer cell can spread to another body part. Therefore, its involvement is the most important prognostic factor in breast cancer patients [52], as it determines the status of breast cancer, possible recurrence, survival, and possible treatment [44, 19].

The SLNB is surgical procedure used to determine if cancer has spread outside the main tumor into the lymphatic system. The fist step is localize the node, for this, the surgeon injects a radioactive substance with the proposal to highlight the nodes are dyed blue. Once the sentinel lymph node is found, the surgeon makes a small incision in the overlying skin and removes the node. Finally, a pathologist then analyzes the sentinel lymph node to check for the presence of cancer cells [59]. In general, sentinel lymph node biopsy is done at the same time the primary tumor is removed, in some cases, it is also possible to do the procedure before or after the tumor is removed. This will depend on how much the lymph ducts have been affected.

The micro-metastasis (MM) are defined small deposits of tumor cells less than 2 mm in diameter. The MM are distinguished from macro-metastasis because they do not have their own source of blood supply and are fed through passive diffusion of nutrients and oxygen,

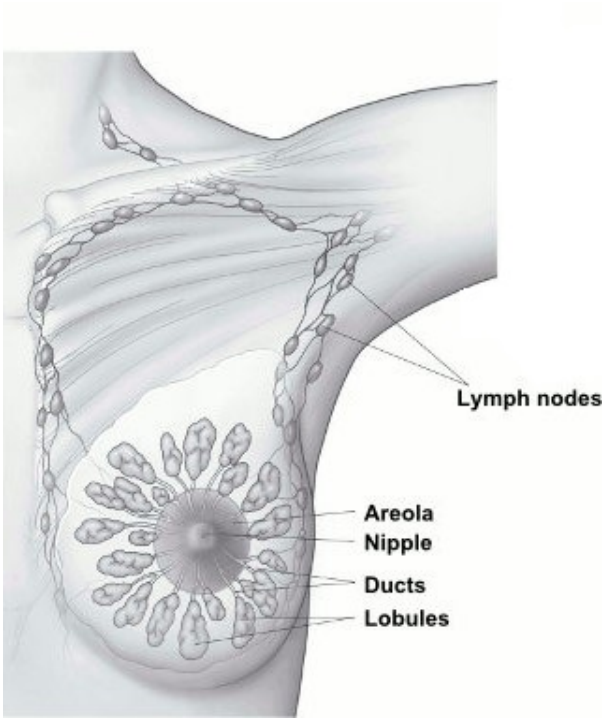


Figure 1-2: Lymph Nodes on the Breast

taken from: www.cancer.org

which limits their growth. These small groups of cells could remain for long periods of time, until the immune system eliminates them or until the formation of new blood vessels through angiogenesis, making the growth of micro-metastasis possible [34]. These small cellular focus are not detected by conventional imaging techniques, requiring the application of more sophisticated methods such as cytometric or molecular imaging.

1.3 Automatic methods to analysis metastasis in the SLN

Every year the number of diagnosed cancer cases increases, taking into account the statistics [47], which implies a high flow of samples to be processed and an increase in the complexity of the cases; due to this, the responsibility of the pathologists is greater. Many computational strategies have been developed to support the work of pathologists, with the proposal to reduce the time consuming and improve the precision in the diagnosis [36, 10, 15, 45]. Digital pathology has the potential to change the way cancer is diagnosed once the imaging in pathology evolves to whole-slide imaging (WSI); these are a great source of information due to high resolution, presence of color information (H&E and immunohistochemistry), and availability of information at multiple scales which allow to serves as an enabling platform for the application of artificial intelligence (AL) in digital pathology [40].

Different computational strategies were created to evaluate the metastasis presence; the principal challenge to evaluate computational systems for the automated detection of metastatic breast cancer in H&E whole slide images of the sentinel lymph node was organized by the International Symposium on Biomedical Imaging (ISBI). Camelyon 16 is the first challenge with the proposal to evaluate the capacity of the algorithm to identify metastasis in the SLN, deep learning algorithms submitted as part of a challenge competition or pathologist interpretation. The results show that the area under the receiver operating characteristic curve (AUC) for the algorithms ranged from 0.556 to 0.994; hence, the algorithm performance was comparable with an expert pathologist interpreting whole-slide images without time constraints [9]. The second challenge is Camelyon 17, the proposal to evaluate algorithms to classify three types of sizes the metastasis (macro-metastasis, micro-metastasis, and Isolated tumor cell) present in the SLN; the best result was combined algorithms of the participants obtaining kappa metric with 0.93. However, the confusion matrix analysis revealed that all participants struggled with reliably identifying the minor type of metastasis, with detection rates below 40% [7, 25]. Studies to detect micro-metastasis include different techniques to identify, as presented by [62]. The principal objective is to identify MM in draining lymph nodes by flow cytometry. Another study report was presented by [11]; the objective is the same however, the method to evaluate was infrared micro-spectral imaging, take into account that the studies on micro-metastasis detect few, exist the need to explore and improve these tasks, due to, the micro-metastasis presence is an essential factor to the determine the prognosis disease and the possible recurrence [6, 55].

1.4 Contribution and Academic products

The detection of micro-metastasis in the sentinel node is a very hard diagnosis and therefore a time consuming process. Yet micro-metastasis is associated with cancer recurrence [57, 55, 58], their precise role and physio-pathology in cancer evolution is still under discussion [18]. They may go undetected in routine exploration by their relatively little occurrence, a crucial limitation to any modern machine learning algorithm. Unlike other approaches, this work proposes a set of robust histopathological color, texture and geometrical descriptors that capture differences in Hematoxylin and Eosin (H&E) stains .

Academic products

Results of this work were published in:

- **Molano, Leidy T.**, Ricardo A. Moncayo, and Eduardo Romero. "Detecting micro-metastases in the sentinel lymph node by characterizing micro-environments." 17th International Symposium on Medical Information Processing and Analysis. Vol. 12088. SPIE, 2021.

1.5 Document Organization

The remaining chapters of this thesis are organized as follows:

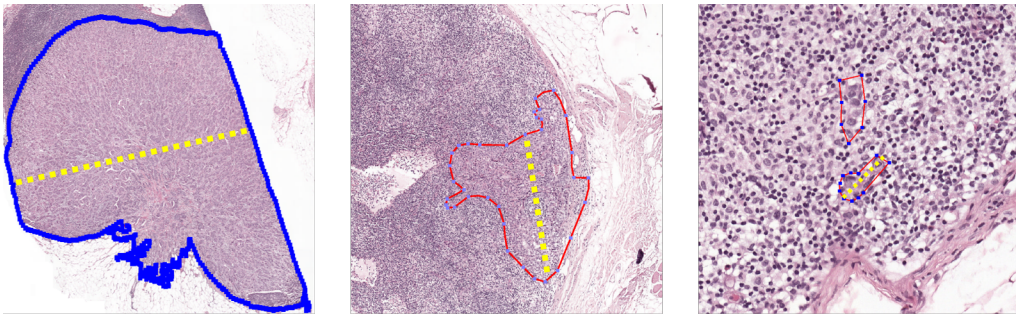
- ***Chapter 2: Detecting micro-metastasis in the sentinel lymph node by characterizing micro-environments***. This chapter presents a classification approach for micro-metastasis detect based on pleomorphic features, as, area, color and texture, these were computed to build feature descriptor.
- ***Chapter 3: Complementary results from experimental setup***. This chapter present the classification task to each pleomorphic feature as, area, color, texture and combination between them. The proposal is to analyze the contribution for each feature independently to identify micro-metatasis.
- ***Chapter4: Conclusions and perspectives***. This final chapter presents the main conclusions of the proposed work, highlighting the main contribution achieved and its impact in the research area. In addition, it depicts some perspectives promoted by this thesis.

2 Detecting micro-metastasis in the sentinel lymph node by characterizing micro-environments

This chapter presents an approach to classify micro-metastasis in H&E images based on an analysis of color area and texture features. For them, the pixel area was calculated from nuclei segmentation, then, a color analysis was performed based on a color space transformation $(R-B, G)$ and finally a texture analysis was performed where five different neighborhoods were analyzed and haralick's texture descriptor were extracted from each one, with this information a feature vector was created to finally classify using an AdaBoost and a SVM with linear kernel. The results obtained an accuracy of 0.86 and an F-score of 0.87 for the proposed model, however, this model was compared with a ResNet 50, obtaining results of 0.74 in accuracy and 0.77 in F-score. The results demonstrated that the proposed model is a robust method to identify the presence of micro-metastasis in H&E images despite the amount of data. A complete version of this chapter has been accepted for publication as a research article in the proceedings of **17th International Seminar on Medical Information Processing and Analysis** (see reference [39]).

2.1 Introduction

Breast cancer (BC) is a neoplasm with approximately 49.7 cases per $100,000$ habitants [12]. Analysis of axillary lymph nodes (ALN) is an important predictor of survival in patients with breast cancer (BC) [27, 17, 60]. Evidence suggests a 3.6% of the patients with axillary node findings have relapses [13]. Node analysis provides information about relapse risk, helps to estimate the risk and treatment [44, 19].The Sentinel Lymph Node (SNL) is the first lymph node reached by metastasising cells from a primary tumour [48], the examination of the SLN can be either positive or negative, this analysis accounts the presence of tumor cells and their respective size to specify the state of cancer and the possible treatment. A positive node can have macro-metastases if cellular clusters are larger than $2\,mm$, micro-metastases when tumor zones vary between $0.2\,mm$ and $2\,mm$ and finally the isolated tumor cells, with sizes smaller than $0.2\,mm$ [49, 25], as illustrated in figure **2-1** Most tumor therapies require metastasis information to determine patient's management and risk.



Diameter: 4mm
Magnification: 1x

(a) Macro-metastasis

Diameter: 0,8mm
Magnification: 10x

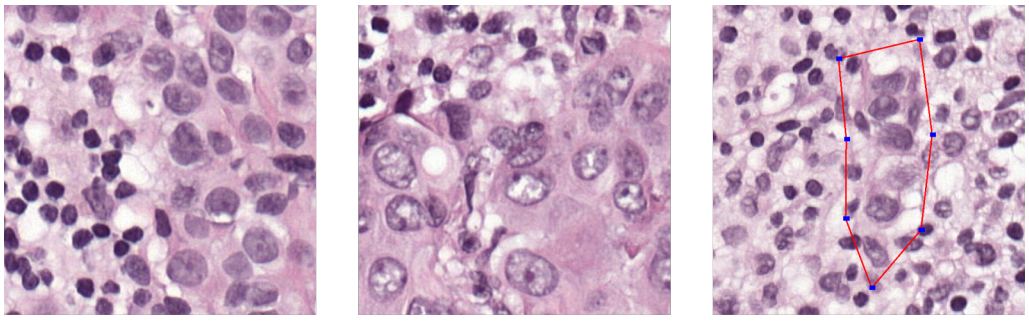
(b) Micro-metastasis

Diameter: 0,05mm
Magnification: 20x

(c) ITC

Figure 2-1: Stratified Metastasis [25]

Overall, micro-metastasis (MM) are missed and the node analysis is usually negative SLN, indicating low recurrence risk, yet patient with micro-metastasis detected have been reported as high risk [17, 44, 30].
Of course the ability to locate micro-metastasis is dependent on the pathologist expertise and the sections from the sample, both factors highly variable and dependent on the reduced size of the affected area[58, 25, 33]. Figure **2-2** shows, at the same scale, how the abnormal nuclei of each tumor stage looks like. The risk introduced by missing micro-metastasis is still unclear, but the risk of missing metastases varies between 2% to 9%[56]. Some studies have reported macro-metastasis detection in SNL is 15.9% greater than in no sentinel lymph node (NSLN), while occult metastases have been identified in 4.09% of SNL and in 0.35% of NSLN [57].



(a) Macro-metastasis

(b) Micro-metástasis

(c) Isoled Tumor Cell

Figure 2-2: Stratified Metastasis [25]. The magnification from these are 40x.

Computational strategies have been developed to identify metastasis, as in the challenge Camelyon 2016 and 2017 in H&E, but the presented investigation is focused in only micro-metastasis detection since this is the actual pathology challenge and the clinic interest. Some studies report pathologists assisted by algorithms improve their report detecting metastasis,

compared to algorithms alone or pathologists alone. Assistance with algorithms has been reported to increase sensitivity of metastasis detection (91% vs. 83%, $p = 0.02$), while the analysis time decreased significantly both for detection of micro-metastasis (61 vs. 116 s, $p = 0.02$) and for negative cases (111 vs. $137s, p = 0.018$)[46]. Most computational approaches have applied deep learning strategies to identify metastasis [9, 54], particularly during the Camelyon challenge organized for the International Symposium on Biomedical Imaging (ISBI), a test for computational systems to automatically detect metastatic breast cancer in whole slide images of the sentinel lymph node. In 2017 the same challenge separated metastasis and micro-metastasis tasks and while detection of metastasis reached an accuracy of 0.91, the micro-metastasis constituted a much more difficult task and the reported mean accuracy was 0.71, being the lowest 0.53 and the highest 0.84 [7]. Few micro-metastasis detection works since then have been reported, but the presented by [11], in which infrared microscopy measures cell populations and micro-metastasis presence is determined by statistic inference.

Clearly then detection of micro-metastasis in the sentinel node is a very hard diagnosis and therefore a time consuming process. Yet micro-metastasis is associated with cancer recurrence [57, 55, 58], their precise role and physio-pathology in cancer evolution is still under discussion [18]. They may go undetected in routine exploration by their relatively little occurrence, a crucial limitation to any modern machine learning algorithm. Unlike other approaches, this work proposes a set of robust histopathological color, texture and geometrical descriptors that capture differences in Hematoxylin and Eosin stains (H&E), nuclei texture at different scales, and shape and size nuclei variations, facilitating interpretation and integration of this analysis with the pathology report. The proposed model is robust and very competitive with CNN approaches to characterize small differences in micro-metastasis.

2.2 Methodology

A main objective of this work was to achieve automatic detection of micro-metastasis using features with meaning to the expert. For so doing, characteristics such as area, color and texture nuclei were captured.

The pipeline of the proposed methodology is presented in figure **2-3**. The method starts by a conventional nuclei detection procedure using a watershed segmentation [16]. Afterwards, lymphocytes are detected, removed and the remaining nuclei segmented to be characterized with area, color, and texture features.

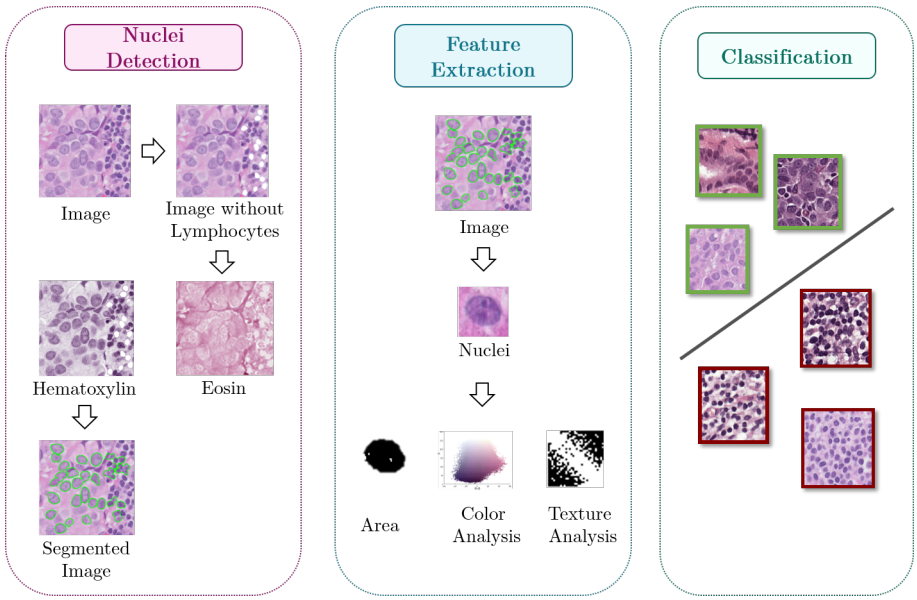


Figure 2-3: Proposed methodology

2.2.1 Color deconvolution and watershed-based nuclei detection

Nuclei are processed in two steps: firstly, lymphocytes are detected and removed from the image with the proposal to improve the abnormal nuclei segmentation, and secondly the remaining cells are segmented by tuning the segmentation watershed algorithm to find tumor cells.

Lymphocytes Segmentation

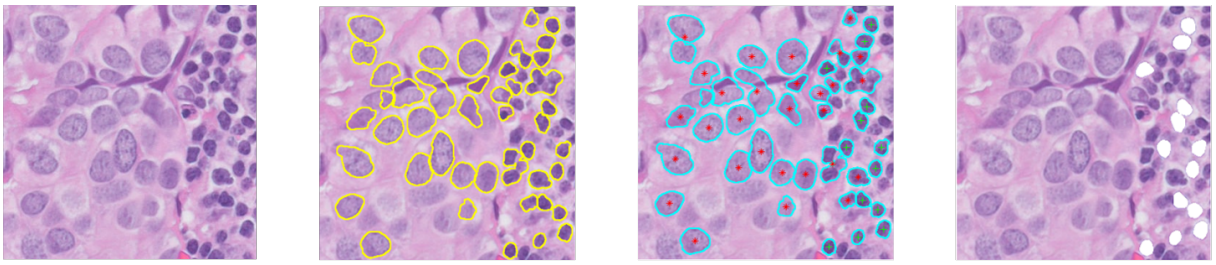
Cells are detected [53] by applying a blob detector watershed [35], as shown in figure **2-1**(b). A support vector machine (SVM) with a polynomial kernel is then trained to differentiate lymphocytes from the other cells using as features: vector area, eccentricity, red color median, entropy and the minimum intensity value as already proposed [16]. In this work, lymphocytes from 4 different cases (206 lymphocytes) were labeled by a pathologist. The model was trained with 140 cells and tested with 66 cells. lymphocyte and non-lymphocyte classes were balanced, the non-lymphocyte class was also examined by the expert. Qualitative results are illustrated in figure **2-4**, and quantitative results shown in table **2-1**. Notice the F-score of 0.86 demonstrates the method acceptably recognizes most lymphocyte patterns.

Results

Model	**Acc**	**Sens**	**Spec**	**F-score**
SVM (Polynomial)	0,86	0.87	0.84	0.86

Table 2-1: Lymphocyte identification training results

This detector is used to find lymphocytes in the complete data-set and remove them from each patch. This lymphocyte removal facilitates the strategy focuses in metastatic nuclei, as shown in the figure **2-4**.

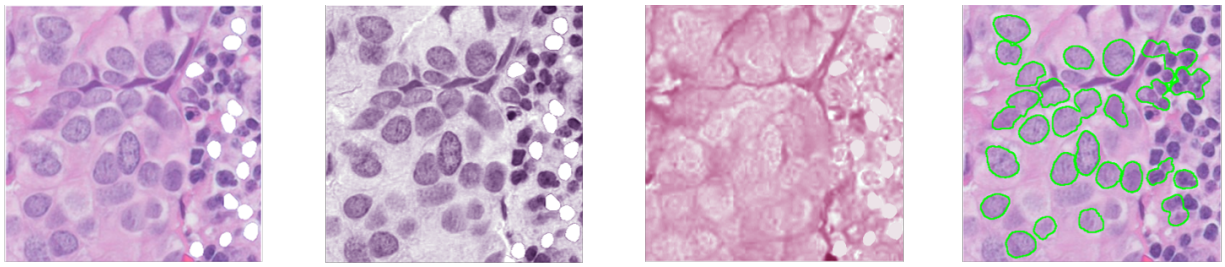


(a) Image (b) Nuclei Segmentation (c) Lymphocyte detection (d) Lymphocyte Removal

Figure 2-4: Process to identify lymphocytes and remove them from the analysis.

Tumoral Nuclei Segmentation

Candidates to Tumor cells are then detected by a re-parameterized watershed algorithm [53] applied to an patch with a minimal number of lymphocytes. Watershed parameters are tuned to detect larger areas in a hematoxylin channel obtained from a color deconvolution technique [37] note that in the staining process the hematoxylin color have the nuclei information. An illustration is presented in figure **2-5**.



(a) Lymphocyte Removal (b) Hematoxylin channel (c) Eosin channel (d) Nuclei Segmentation

Figure 2-5: Process to segment abnormal nuclei

2.2.2 Feature extraction based on area, color and texture.

Nuclei variations (Pleomorphism) influence the type of cells detected. The analysis uses area, color and texture patterns. This information represents a small patch with a variable number of nuclei, reason why each feature vector per patch is normalized by the number of nuclei in the patch.

Area Analysis

Each nucleus is characterized by its area and the patch is represented by the histogram of occurrences of the possible area values in the database. For so doing, the minimum and maximum area values for the whole dataset is split into 10 bins.

Color Analysis

The purpose of this analysis is to compute Hematoxylin and Eosin (H&E) color variations. Basically, normal cells are intensely purple while metastatic ones are smoothly pink, a difference captured by projecting the original (R, G, B) nucleus to the $[(R-B), G]$ plane. The G axis is divided in 25 intervals and each of them is in due turn split into 30 bins in the direction of the projected $R-B$ axis. Each R, G, B pixel from the patch is a occurrence in a unique bin of this bi-dimensional conditional probability. The obtained feature vector is composed of 750 values, i.e., 25 histograms $\times$ 30 bins ,as shown in figure **2-6**.

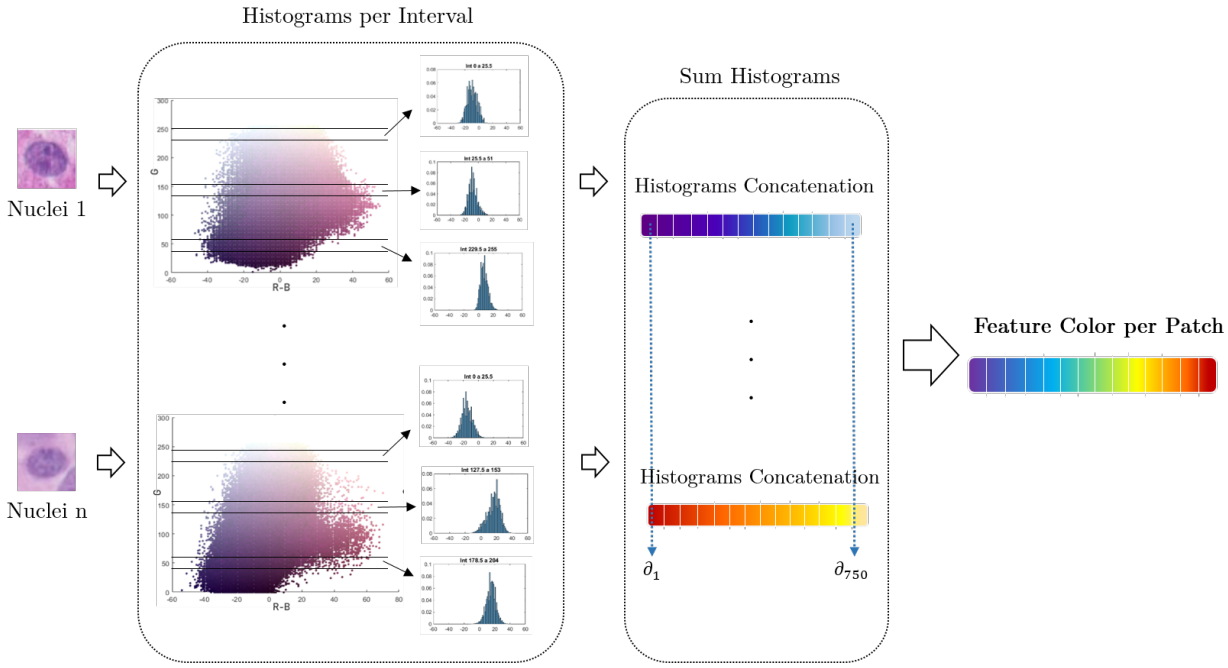


Figure 2-6: Color Analysis

Texture Analysis

A first step of the texture analysis is to select a gray level channel with an optimal contrast between intensity levels. Nuclei images in the primary color space (R, G, B) are transformed to the (L, A, B) color space and the B channel is selected since it is the bluest one.
The texture analysis computes second-order statistical moments from a gray level co-occurrence matrix (GLCM), which esentially is a discrete approximation to a marginal spatial probabil-

ity distribution. GLCM was quantized to 64 gray levels and was calculated for diagonal and vertical-horizontal neighbors and different scales from the first up to the fifth order. Then, the 14 Haralick's texture descriptors are calculated by each the five scales of co-occurrence [26], as illustrated in figure **2-7**. Finally, each Haralick's feature is represented as a histogram of occurrences by quantizing the dynamic range, the difference between the maximum and minimum values for the whole database, into ten intervals. The final vector of frequencies is normalized.

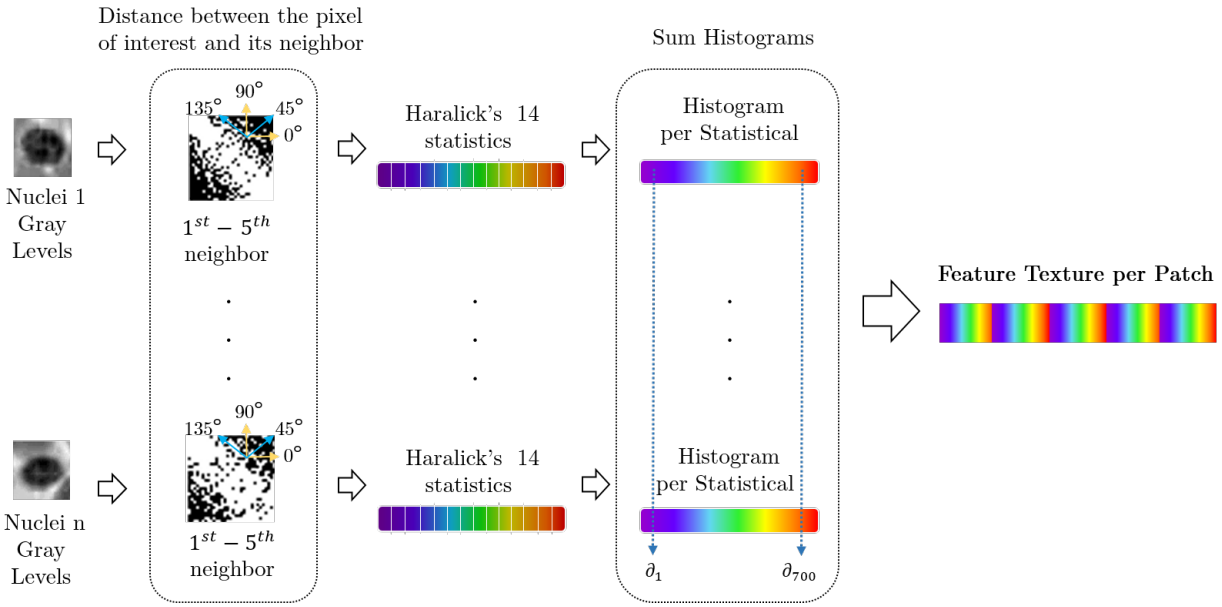


Figure 2-7: Texture analysis

2.2.3 Classification

A binary classification task was performed by detecting patches with high probability of having tumoral nuclei. Positive class was a patch with metastases and negative class corresponded to patches with normal cells. A feature matrix was constructed by concatenating the resulting feature vectors and it was used to train AdaBoost and SVM with lineal kernel models. Performance metrics were accuracy (Acc), sensitivity (Sens), specificity (Spec) and F-score.

2.3 Experimental Setup

The proposed approach was evaluated in the Camelyon 16 Database [25], which contains a total of 50 annotated cases, but only 25 with micro-metastasis and macro-metastasis were used ,the remaining cases were excluded since are labeled either as isolated tumor cells or

low quality WSI. From each case, patches of 350 × 350 pixels were extracted from regions annotated by an expert as containing macro-metastasis, micro-metastasis and normal tissue patches.

A model is built with a training-validate set composed of 18 cases with normal and metastatic tissue, including micro and macro metastasis patches.As a result, a total of 3222 patches were extracted, 1611 with metastases and 1611 control patches. Then, under a 10-fold cross validation the accuracy in each fold is computed and the model with higher accuracy is selected and used to validate the method in a new sets of cases: 5 cases only with micro-metastasis, as total 334 patches to test.

AdaBoost and SVM with lineal kernel classifiers were trained. Both validation sets were also compared with a ResNet50 [28] a convolutional neural network (CNN) trained from scratch, with a learning rate of $1 \times 10^{-}4$, a batch size of 64, with data augmentation and rotations, vertical and Horizontal flips, random brightness, random Contrast and RGB shift with 200 epochs. This CNN model was also evaluated with transfer learning using ImageNet to pre-training weights . The network uses an ADAM optimizer and the loss function is the categorical crossentropy. The best model tested was the one with the lowest value of the loss function during training/validation. The ResNet50 was selected because of its good performance in the detection of metastasis [9].

2.4 Results

The results of the 10-fold validation scheme are shown in table **2-2**. The mean of the 10 folds for ACC, Sens, Spec and F-score are presented for the models trained with AdaBoost and SVM with lineal kernel. The model trained with the AdaBoost (0.79 F-score) classifier outperforms the SVM with lineal kernel(0.73 F-score) for all metrics. These models are applied to evaluate the blind micro-metastasis dataset.

Validation (K-fold)

Model	Acc		Sens		Spec		F-score	
	Mean	Best Value	Mean	Best Value	Mean	Best Value	Mean	Best Value
AdaBoost	0.80(±0.06)	0.86	0.75(±0.01)	0.77	0.85(±0.09)	0.94	0.79(±0.07)	0.84
SVM Lineal	0.76(±0.07)	0.89	0.68(±0.13)	0.81	0.84(±0.10)	0.97	0.73(±0.011)	0.88

Table 2-2: Results of 10-fold cross validation and the best model.

Results of the validation with micro-metastasis are presented in table **2-3**. The method with AdaBoost classifier outperforms the SVM with linear kernel, F-scores were respectively 0.88 and 0.82. Interestingly, the obtained F-score with the ResNet-50 plus data augmentation was 0.77, while this F-score was 0.61 without data augmentation. The confuse matrix show in the table

Results - Test

	Model	Acc	Sens	Spec	F-score
Proposed Approach	AdaBoost	**0.86**	**0.89**	**0.83**	**0.87**
Proposed Approach	SVM Lineal	0.82	0.96	0.69	0.84
Deep Learning - Base without Augmentation	ResNet -50	0.64	0.56	0.73	0.61
Deel Learning - Base Augmentation	ResNet -50	0.74	0.86	0.63	0.77

Table 2-3: table of results for the micro-metastasis dataset.

These results with a standard CNN are similar to the ones reported in the challenge Camelyon 17 to identify micro-metastasis. The four best teams informed accuracies varying from 0.53 to 0.83 and using different deep learning architectures. The presented approach with AdaBoost obtained an accuracy 0.86 for the same task.

2.5 Discussion and Conclusion

This work proposes a method to find out micro-metastasis, very likely the most difficult task when exploring the sentinel node. Micro-metastasis role in the dissemination of cancer is still unclear, probably because they go undetected in many cases. A better understanding of their role pass by detecting them with high confidence levels and this is not the case so far.

The proposed characterization showed similar results to classify macro-metastasis (data not shown), but above all outperformed the other approaches at finding out micro-metastasis. Although deep learning approaches are widely used in many medical applications, the few available micro-metastasis data are a limitation for their application. In contrast, approaches like the one herein presented are less dependent on data, more stable and completely interpretable in terms of the classic pathology analysis.

Future studies comprises the use of more cases and the accurate segmentation of the micro-metastasis in large population studies to estimate patient's risk and prognosis.

3 Evaluation and analysis of proposed features

This chapter is a complementary experimental setup with the proposal to evaluate the performance of the model, the idea is analyze the contribution of each feature from separate (area,color and texture).

3.1 Complementary Results

The proposal approach was evaluated each feature for separate with the proposal to analyze the importance of each feature extracted to identify the micro-metastasis. The experimental set up is similar to the last analysis, however, the difference was evaluated each feature for separated. The experimental setup is the same to the section (2.3), the evaluation was in the Camelyon 16 Database [25], the distribution of the cases is the same. A model from each feature is built with a training-validate set composed of 18 cases with normal and metastatic tissue, including micro and macro metastasis patches. Then, under a 10-fold cross validation the accuracy in each fold is computed. So, the fold where the model select in the previous analysis with highlight accuracy is consider to the select each model to analyzed each feature with the proposal to does not have variability in the experiment.Then, each model is validate in the blind set with the same 5 cases. The model to trainer are AdaBoost and SVM with linear kernel to area, color and texture feature.

3.1.1 Results

The results of the 10-fold validation scheme for each feature are shown to continue. The mean on the 10 folds for ACC,Sens,Spec and F-score are presented for each model trained with AdaBoost and SVM with linear kernel.

Once we have the major model, an additional experiment was selected depending best performance to identify micro-metastasis. In this case, the color feature was combined with the area and texture to analyze the models.

Results to the Area Feature

The results of the 10-fold validation scheme to the area feature are show in the table **3-1**. The model training with AdaBoost (0.68 F-score) outperforms the SVM with linear kernel (0.66 F-score), however, the Specificity for both models is higher.

Validation (K-fold)

Model	Acc		Sens		Spec		F-score	
	Mean	Best Value	Mean	Best Value	Mean	Best Value	Mean	Best Value
AdaBoost	0.70(±0.07)	0.78	0.66(±0.16)	0.82	0.74(±0.06)	0.74	0.68(±0.10)	0.78
SVM Lineal	0.71(±0.09)	0.82	0.59(±0.18)	0.77	0.82(±0.07)	0.87	0.66(±0.013)	0.81

Table 3-1: Results of 10-fold cross validation and the best model for the Area feature.

The results for the micro-metastasis from area feature independent are shown in the table **3-2**. The results evidence that the AdaBoost classifier (0.78 F-score) outperforms the SVM with the linear kernel(0.75 F-score). Although both models have the same accuracy (0.77 Acc), the best model to identify the true positive is the AdaBoost (0.80 Sens); this means that it is the best model for detecting cases with the true presence of micro-metastasis.

Results - Test

Model	Acc	Sens	Spec	F-score
AdaBoost	**0.77**	**0.80**	**0.73**	**0.78**
SVM Lineal	0.77	0.71	0.83	0.75

Table 3-2: Results of micro-metastasis dataset with Area feature only.

Results to the Color Feature

The results of the 10-fold validation scheme to the color feature are show in the table **3-3**. The model training with AdaBoost (0.78 F-score) outperforms the SVM with linear kernel (0.73 F-score) in all metrics.

Validation (K-fold)

Model	Acc		Sens		Spec		F-score	
	Mean	Best Value	Mean	Best Value	Mean	Best Value	Mean	Best Value
AdaBoost	0.79(±0.05)	0.82	0.76(±0.13)	0.71	0.81(±0.09)	0.93	0.78(±0.06)	0.81
SVM Lineal	0.75(±0.07)	0.78	0.72(±0.21)	0.79	0.78(±0.17)	0.77	0.73(±0.11)	0.79

Table 3-3: Results of 10-fold cross validation and the best model for the Color feature.

The results for the micro-metastasis from color feature independent are shown in the table **3-4**. The results evidence that each feature contributes to the detection; the color feature is the most relevant feature, taking into account the model with SVM with the linear kernel (0.87 F-score) outperforms the AdaBoost(.82 F-score). The SVM classification is a model

identify micro-metastasis with an accuracy highlight (0.87); however, the model has high specificity (0.87) to detect the true negative cases.

Results - Test

Model	Acc	Sens	Spec	F-score
AdaBoost	0.82	0.84	0.80	0.82
SVM Lineal	**0.87**	**0.86**	**0.87**	**0.87**

Table 3-4: Results of micro-metastasis dataset with Color feature only.

Results to the texture Feature

The results of the 10-fold validation scheme to the texture feature are show in the table **3-5**. The model training with SVM with linear kernel (0.76 F-score) outperforms the AdaBoost(0.70 F-score) in all metrics.

Validation (K-fold)

Model	Acc		Sens		Spec		F-score	
	Mean	Best Value	Mean	Best Value	Mean	Best Value	Mean	Best Value
AdaBoost	0.71(±0.05)	0.74	0.70(±0.15)	0.78	0.71(±0.12)	0.70	0.70(±0.07)	0.75
SVM Lineal	0.75(±0.07)	0.83	0.79(±0.14)	0.79	0.72(±0.01)	0.88	0.76(±0.08)	0.83

Table 3-5: Results of 10-fold cross validation and the best model for the Texture feature.

The results for the micro-metastasis from texture feature independent are shown in the table **3-6**. The results evidence that the micro-metastasis detection with texture feature is performed with the AdaBoost model (0.84 F-score) compared with the SVM with the linear kernel (0.80 F-score). This model has a high probability of detecting true negative cases (0.84 Spec).

Results - Test

Model	Acc	Sens	Spec	F-score
AdaBoost	**0.84**	**0.83**	**0.84**	**0.84**
SVM Lineal	0.77	0.92	0.62	0.80

Table 3-6: Results of micro-metastasis dataset with Texture feature only.

Once the model with the best contribution to identify micro-metastasis was defined, in this case, the color characteristics, we combined color-area and color-texture. The results are shown below:

Results to the Color and Area Feature

Once The results of the 10-fold validation scheme to the color and area features are show in the table **3-7**. The model training with AdaBoost (0.79 F-score) outperforms the SVM with linear kernel (0.77 F-score) in all metrics.

Validation (K-fold)

Model	Acc		Sens		Spec		F-score	
	Mean	Best Value	Mean	Best Value	Mean	Best Value	Mean	Best Value
AdaBoost	0.80(±0.06)	0.86	0.78(±0.14)	0.81	0.82(±0.07)	0.91	0.79(±0.07)	0.86
SVM Lineal	0.79(±0.06)	0.81	0.75(±0.16)	0.68	0.83(±0.12)	0.94	0.77(±0.08)	0.78

Table 3-7: Results of 10-fold cross validation and the best model for the color and area feature.

The micro-metastasis results from the color-area feature are shown in the table **3-6**. The results evidence that each feature contributes to the detection; the color feature is the most relevant.The SVM with the linear kernel model (0.89 F-score) outperforms the AdaBoost (0.87 F-score). However, the SVM model has a high probability of detecting true negative cases (0.92 Spec) compared with AdaBoost (0.84 Spec). Considering that the idea is to find true micro-metastasis, the model AdaBoost has better results in Sensitivity (0.89 Sens) than the SVM (0.84 Sens).

Results - Test

Model	Acc	Sens	Spec	F-score
AdaBoost	0.86	0.89	0.84	0.87
SVM Lineal	**0.89**	**0.87**	**0.92**	**0.89**

Table 3-8: Results of micro-metastasis dataset with color and area features.

Results to the Color and Texture Feature

The results of the 10-fold validation scheme to the color and texture feature are show in the table **3-9**. The model training with AdaBoost (0.76 F-score) outperforms the SVM with linrar kernel (0.73 F-score) in all metrics.

Validation (K-fold)

Model	Acc		Sens		Spec		F-score	
	Mean	Best Value	Mean	Best Value	Mean	Best Value	Mean	Best Value
AdaBoost	0.77(±0.01)	0.79	0.73(±0.02)	0.66	0.81(±0.01)	0.92	0.76(±0.01)	0.76
SVM Lineal	0.75(±0.02)	0.83	0.72(±0.03)	0.79	0.79(±0.009)	0.87	0.73(±0.03)	0.83

Table 3-9: Results of 10-fold cross validation and the best model for the color and texture feature.

The results for the micro-metastasis from each feature independent are shown in the table **3-6**. The results evidence the model with SVM with the linear kernel (0.86 F-score) outperforms the AdaBoost (0.85 F-score); this model has a high probability of detecting true positive cases (0.96 Sens); however, the accuracy is slower compared as this.

Results - Test

Model	Acc	Sens	Spec	F-score
AdaBoost	0.85	0.86	0.83	0.85
SVM Lineal	0.84	0.96	0.72	0.86

Table 3-10: Results of micro-metastasis dataset with color and texture features.

3.1.2 Discussion and Conclusion

This step to analyze the proposal approach evidence that each method supports the micro-metastasis detection, and each feature show significant provides that improve the detection. The color methodology is the approach that contributes the most to the identification of micro-metastasis. The combination between color and area demonstrates that it is possible to identify micro-metastasis using only this. Last but not least, this study comprises more cases and explores the proposal technique in other cases to identify if, only using the color and/or area, it is possible to characterize cancerous cells to estimate the patient with risk.

4 Conclusions and perspectives

4.1 Conclusions

This thesis has developed an approach to detect micro-metastasis from basic characteristics such as area, color, and nuclei texture through histopathology images stained with hematoxylin and eosin. One of the contributions of this work has been the identification of micro-metastasis through pleomorphic features taking into account the main features used by experts for histopathological analysis.

This work demonstrated that color analysis could become a more robust descriptor in identifying cancer cells, taking into account that color is one of the most informative features in histopathology analysis.

We were able to introduce a descriptor to characterize histopathology images of SLNs by capturing area, color, and texture features that proved to be useful for the identification of micro-metastasis despite the small amount of data, in contrast to the results obtained from neural networks where it is evident that the lack of data influences the low performance for the detection of micro-metastasis.

4.2 Perspectives

This thesis has contributed to identifying micro-metastasis from the features of area, color, and texture in the sentinel lymph node, which are pleomorphic characteristics frequently used by pathologists in histopathological analysis. Despite the quantity of data evaluated, the model presents a good performance in contrast to neural networks, which have been characterized by requiring a large quantity of data and higher computational costs to present better results. This proposed methodology is a door that, in the future, can be implemented in another histopathological problem that requires the analysis of nuclei due to the similarity of characteristics analyzed by the pathologists. Future work will evaluate the model on another data set to assess feasibility and performance without forgetting the possibility of the preceding recurrence of lymph node analysis using part of the proposed model.

References

[1] *Cancer Staging — cancer.org.* https://www.cancer.org/treatment/understanding-your-diagnosis/staging.html. – [Accessed 30-Nov-2022]

[2] *CxE1;ncer de mama — CancerQuest — cancerquest.org.* https://www.cancerquest.org/es/para-los-pacientes/cancer-por-tipo/cancer-de-mama. – [Accessed 01-Dec-2022]

[3] *CÃ¡ncer de mama - Estadios — cancer.net.* https://www.cancer.net/es/tipos-de-cancer-de-mama/estadios. – [Accessed 30-Nov-2022]

[4] *Types of Breast Cancer — About Breast Cancer — cancer.org.* https://www.cancer.org/cancer/breast-cancer/about/types-of-breast-cancer.html. – [Accessed 30-Nov-2022]

[5] *What Is Breast Cancer? — American Cancer Society — cancer.org.* https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html. – [Accessed 30-Nov-2022]

[6] Apple, Sophia K.: Sentinel lymph node in breast cancer: review article from a pathologistâ€™s point of view. En: *Journal of pathology and translational medicine* 50 (2016), Nr. 2, p. 83

[7] Bandi, Peter ; Geessink, Oscar ; Manson, Quirine ; Van Dijk, Marcory ; Balkenhol, Maschenka ; Hermsen, Meyke ; Bejnordi, Babak E. ; Lee, Byungjae ; Paeng, Kyunghyun ; Zhong, Aoxiao [u. a.]: From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. En: *IEEE transactions on medical imaging* 38 (2018), Nr. 2, p. 550–560

[8] Bass, Siddharth S. ; Cox, Charles E. ; Ku, Ni N. ; Berman, Claudia ; Reintgen, Douglas S.: The role of sentinel lymph node biopsy in breast cancer. En: *Journal of the American College of Surgeons* 189 (1999), Nr. 2, p. 183–194

[9] Bejnordi, Babak E. ; Veta, Mitko ; Van Diest, Paul J. ; Van Ginneken, Bram ; Karssemeijer, Nico ; Litjens, Geert ; Van Der Laak, Jeroen A. ; Hermsen, Meyke ; Manson, Quirine F. ; Balkenhol, Maschenka [u. a.]: Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. En: *Jama* 318 (2017), Nr. 22, p. 2199–2210

[10] Bera, Kaustav ; Schalper, Kurt A. ; Rimm, David L. ; Velcheti, Vamsidhar ; Madabhushi, Anant: Artificial intelligence in digital pathologyâ€"new tools for diagnosis and precision oncology. En: *Nature reviews Clinical oncology* 16 (2019), Nr. 11, p. 703–715

[11] Bird, Benjamin ; Bedrossian, Kristi ; Laver, Nora ; Miljković, Miloš ; Romeo, Melissa J. ; Diem, Max: Detection of breast micro-metastases in axillary lymph nodes by infrared micro-spectral imaging. En: *Analyst* 134 (2009), Nr. 6, p. 1067–1076

[12] Bray, Freddie ; Ferlay, Jacques ; Soerjomataram, Isabelle ; Siegel, Rebecca L. ; Torre, Lindsey A. ; Jemal, Ahmedin: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. En: *CA: a cancer journal for clinicians* 68 (2018), Nr. 6, p. 394–424

[13] Castro, CL ; Amat, M ; Mysler, DC ; Costanzo, MV ; Nervo, AA ; Coló, FA ; Loza, JB ; Chacón, R: Ganglio centinela positivo sin linfadenectomía axilar: aplicabilidad de las nuevas guías internacionales. En: *REVISTA ARGENTINA DE MASTOLOGíA* 36 (2017), Nr. 130

[14] Chen, Jia-Mei ; Li, Yan ; Xu, Jun ; Gong, Lei ; Wang, Lin-Wei ; Liu, Wen-Lou ; Liu, Juan: Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: A review. En: *Tumor Biology* 39 (2017), Nr. 3, p. 1010428317694550

[15] Colling, Richard ; Pitman, Helen ; Oien, Karin ; Rajpoot, Nasir ; Macklin, Philip ; in Histopathology Working Group, CM-Path A. ; Bachtiar, Velicia ; Booth, Richard ; Bryant, Alyson ; Bull, Joshua [u. a.]: Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. En: *The Journal of pathology* 249 (2019), Nr. 2, p. 143–150

[16] Corredor, Germán ; Wang, Xiangxue ; Lu, Cheng ; Velcheti, Vamsidhar ; Romero, Eduardo ; Madabhushi, Anant: A watershed and feature-based approach for automated detection of lymphocytes on lung cancer images. En: *Medical Imaging 2018: Digital Pathology* Vol. 10581 International Society for Optics and Photonics, 2018, p. 105810R

[17] Cox, Charles E. ; Kiluk, John V. ; Riker, Adam I. ; Cox, John M. ; Allred, Nathon ; Ramos, Daniel C. ; Dupont, Elisabeth L. ; Vrcel, Vesna ; Diaz, Nils ; Boulware, David: Significance of sentinel lymph node micrometastases in human breast cancer. En: *Journal of the American College of Surgeons* 206 (2008), Nr. 2, p. 261–268

[18] Cserni, G ; Amendoeira, I ; Apostolikas, N ; Bellocq, JP ; Bianchi, S ; Bussolati, Giovanni ; Boecker, W ; Borisch, B ; Connolly, CE ; Decker, T [u. a.]:

Pathological work-up of sentinel lymph nodes in breast cancer. Review of current data to be considered for the formulation of guidelines. En: *European Journal of Cancer* 39 (2003), Nr. 12, p. 1654–1667

[19] Cserni, Gábor ; Bianchi, Simonetta ; Boecker, Werner ; Decker, Thomas ; Lacerda, Manuela ; Rank, Fritz ; Wells, Clive A. ; for Breast Screening Pathology, European Working G.: Improving the reproducibility of diagnosing micrometastases and isolated tumor cells. En: *Cancer* 103 (2005), Nr. 2, p. 358–367

[20] Danko, Melissa E. ; Bennett, Kyla M. ; Zhai, Jun ; Marks, Jeffrey R. ; Olson Jr, John A.: Improved staging in node-positive breast cancer patients using lymph node ratio: results in 1,788 patients with long-term follow-up. En: *Journal of the American College of Surgeons* 210 (2010), Nr. 5, p. 797–805

[21] for Disease Control, Centers ; Prevention [u. a.]. *Breast cancer in young women https://www. cdc. gov/cancer/breast/young_women/bringyourbrave/pdf/breastcanceryoungwomenfactsheet. pdf*. 2018

[22] Duncan, W ; Kerr, GR: The curability of breast cancer. En: *Br Med J* 2 (1976), Nr. 6039, p. 781–783

[23] Giammarile, Francesco ; Vidal-Sicart, Sergi ; Paez, Diana ; Pellet, Olivier ; Enrique, Estrada-Lobato ; Mikhail-Lette, Miriam ; Morozova, Olga ; Camila, Navarro Marulanda M. ; Ivonne, Rodríguez Sanchez Diana ; Bolton, Roberto C D. [u. a.]: Sentinel Lymph Node Methods in Breast Cancer. En: *Seminars in Nuclear Medicine* Elsevier, 2022

[24] Giuliano, Armando E. ; Dale, Paul S. ; Turner, Roderick R. ; Morton, Donald L. ; Evans, Sheila W. ; Krasne, David L.: Improved axillary staging of breast cancer with sentinel lymphadenectomy. En: *Annals of surgery* 222 (1995), Nr. 3, p. 394

[25] Halilovic, Altuna ; EhteshamiÂ Bejnordi, Babak ; vanÂ derÂ Laak, Jeroen ; Balkenhol, Maschenka ; Hermsen, Meyke ; Geessink, Oscar ; Bult, Peter ; vanÂ deÂ Loo, Rob ; Vogels, Rob ; Litjens, Geert ; Stathonikos, Nikolas ; vanÂ Diest, Paul ; Bandi, Peter ; Manson, Quirine F. ; Baidoshvili, Alexi ; Wauters, Carla ; vanÂ Dijk, Marcory: 1399 H & amp;E-stained sentinel lymph node sections of breast cancer patients: the CAMELYON dataset. En: *GigaScience* 7 (2018), 05, Nr. 6. – ISSN 2047–217X

[26] Haralick, Robert M. ; Shanmugam, Karthikeyan ; Dinstein, Its' H.: Textural features for image classification. En: *IEEE Transactions on systems, man, and cybernetics* (1973), Nr. 6, p. 610–621

[27] Hatoum, Hassan A. ; Jamali, Faek R. ; El-Saghir, Nagi S. ; Musallam, Khaled M. ; Seoud, Muhieddine ; Dimassi, Hani ; Abbas, Jaber ; Khalife, Mohamad ; Boulos, Fouad I. ; Tawil, Ayman N. [u. a.]: Ratio between positive lymph nodes and total excised axillary lymph nodes as an independent prognostic factor for overall survival in patients with nonmetastatic lymph node-positive breast cancer. En: *Indian journal of surgical oncology* 1 (2010), Nr. 1, p. 68–75

[28] He, Kaiming ; Zhang, Xiangyu ; Ren, Shaoqing ; Sun, Jian: Deep residual learning for image recognition. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, p. 770–778

[29] Herrán, Oscar F. ; Álvarez, Diana C. ; Quintero-Lesmes, Doris C.: Dietary patterns and breast cancer in Colombia: an ecological study. En: *International health* 12 (2020), Nr. 4, p. 317–324

[30] Houvenaeghel, G ; Sabatier, R ; Reyal, F ; Classe, JM ; Giard, S ; Charitansky, H ; Rouzier, R ; Faure, C ; Garbay, JR ; Daraï, E [u. a.]: Axillary lymph node micrometastases decrease triple-negative early breast cancer survival. En: *British journal of cancer* 115 (2016), Nr. 9, p. 1024–1031

[31] IDIAZABAL, DRA GABRIELA M.: Histología del sistema linfático. En: *Flebología y Linfología* 13 (2010), p. 797–804

[32] Jarkman, Sofia ; Karlberg, Micael ; Pocevičiūtė, Milda ; Bodén, Anna ; Bándi, Péter ; Litjens, Geert ; Lundström, Claes ; Treanor, Darren ; van der Laak, Jeroen: Generalization of Deep Learning in Digital Pathology: Experience in Breast Cancer Metastasis Detection. En: *Cancers* 14 (2022), Nr. 21, p. 5424

[33] Bouquet de Jolinière, Jean ; Major, A ; Khomsi, F ; Ben Ali, N ; Guillou, L ; Feki, A: The Sentinel Lymph Node in Breast Cancer: Problems Posed by Examination During Surgery. A Review of Current Literature and Management. En: *Frontiers in Surgery* 5 (2018), p. 56

[34] Kell, MR ; Winter, DC ; O'sullivan, GC ; Shanahan, F ; Redmond, HP: Biological behaviour and clinical implications of micrometastases. En: *British journal of surgery* 87 (2000), Nr. 12, p. 1629–1639

[35] Levner, Ilya ; Zhang, Hong: Classification-driven watershed segmentation. En: *IEEE Transactions on Image Processing* 16 (2007), Nr. 5, p. 1437–1445

[36] Litjens, Geert ; Sánchez, Clara I. ; Timofeeva, Nadya ; Hermsen, Meyke ; Nagtegaal, Iris ; Kovacs, Iringo ; Hulsbergen-Van De Kaa, Christina ; Bult, Peter ; Van Ginneken, Bram ; Van Der Laak, Jeroen: Deep learning as a tool for increased

accuracy and efficiency of histopathological diagnosis. En: *Scientific reports* 6 (2016), Nr. 1, p. 1–11

[37] Macenko, Marc ; Niethammer, Marc ; Marron, James S. ; Borland, David ; Woosley, John T. ; Guan, Xiaojun ; Schmitt, Charles ; Thomas, Nancy E.: A method for normalizing histology slides for quantitative analysis. En: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* IEEE, 2009, p. 1107–1110

[38] Meyer, John S. ; Alvarez, Consuelo ; Milikowski, Clara ; Olson, Neal ; Russo, Irma ; Russo, Jose ; Glass, Andrew ; Zehnbauer, Barbara A. ; Lister, Karen ; Parwaresch, Reza: Breast carcinoma malignancy grading by Bloom–Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. En: *Modern pathology* 18 (2005), Nr. 8, p. 1067–1078

[39] Molano, Leidy T. ; Moncayo, Ricardo A. ; Romero, Eduardo: Detecting micrometastases in the sentinel lymph node by characterizing micro-environments. En: *17th International Symposium on Medical Information Processing and Analysis* Vol. 12088 SPIE, 2021, p. 313–321

[40] Niazi, Muhammad Khalid K. ; Parwani, Anil V. ; Gurcan, Metin N.: Digital pathology and artificial intelligence. En: *The lancet oncology* 20 (2019), Nr. 5, p. e253–e261

[41] Patani, Neill ; Mokbel, Kefah: The clinical significance of sentinel lymph node micrometastasis in breast cancer. En: *Breast cancer research and treatment* 114 (2009), Nr. 3, p. 393–402

[42] Patani, Neill ; Mokbel, Kefah: Clinical significance of sentinel lymph node isolated tumour cells in breast cancer. En: *Breast cancer research and treatment* 127 (2011), Nr. 2, p. 325–334

[43] Rodrigues-Fernandes, Carla I. ; Speight, Paul M. ; Khurram, Syed A. ; Araújo, Anna Luíza D. ; da Cruz Perez, Danyel E. ; Fonseca, Felipe P. ; Lopes, Márcio A. ; de Almeida, Oslei P. ; Vargas, Pablo A. ; Santos-Silva, Alan R.: The use of digital microscopy as a teaching method for human pathology: a systematic review. En: *Virchows Archiv* (2020), p. 1–12

[44] Salhab, Mohamed ; Patani, Neill ; Mokbel, Kefah: Sentinel lymph node micrometastasis in human breast cancer: an update. En: *Surgical oncology* 20 (2011), Nr. 4, p. e195–e206

[45] Salto-Tellez, Manuel ; Maxwell, Perry ; Hamilton, Peter: Artificial intelligence-the third revolution in pathology. En: *Histopathology* 74 (2019), Nr. 3, p. 372–376

[46] Steiner, David F. ; MacDonald, Robert ; Liu, Yun ; Truszkowski, Peter ; Hipp, Jason D. ; Gammage, Christopher ; Thng, Florence ; Peng, Lily ; Stumpe, Martin C.: Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. En: *The American journal of surgical pathology* 42 (2018), Nr. 12, p. 1636

[47] Sung, Hyuna ; Ferlay, Jacques ; Siegel, Rebecca L. ; Laversanne, Mathieu ; Soerjomataram, Isabelle ; Jemal, Ahmedin ; Bray, Freddie: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. En: *CA: a cancer journal for clinicians* 71 (2021), Nr. 3, p. 209–249

[48] Tanis, Pieter J. ; Nieweg, Omgo E. ; Olmos, Renato A V. ; Emiel, J ; Kroon, Bin B.: History of sentinel node and validation of the technique. En: *Breast Cancer Research* 3 (2001), Nr. 2, p. 1–4

[49] Tavassoli, Fattaneh A.: World Health Organization classification of tumors. En: *Pathology and genetics of tumors of the breast and female genital organs* 97 (2003)

[50] Tjan-Heijnen, Vivianne C. ; Bult, Peter ; de Widt-Levert, Louise M. ; Ruers, Theo J. ; Beex, Louk V.: Micro-metastases in axillary lymph nodes: an increasing classification and treatment dilemma in breast cancer due to the introduction of the sentinel lymph node procedure. En: *Breast cancer research and treatment* 70 (2001), Nr. 2, p. 81–88

[51] Tonellotto, Fabiana ; Bergmann, Anke ; de Souza Abrahão, Karen ; de Aguiar, Suzana S. ; Bello, Marcelo A. ; Thuler, Luiz Claudio S.: Impact of number of positive lymph nodes and lymph node ratio on survival of women with node-positive breast cancer. En: *European journal of breast health* 15 (2019), Nr. 2, p. 76

[52] Tresserra, F ; Grases, PJ ; Izquierdo, M ; Fabregas, R ; Fernandez-Cid, A: Detection of micrometastases in the sentinel node of patients with infiltrating breast cancer. En: *PROGRESOS DE OBSTETRICIA Y GINECOLOGIA* 45 (2002), Nr. 2, p. 37–44

[53] Veta, Mitko ; Van Diest, Paul J. ; Kornegoor, Robert ; Huisman, André ; Viergever, Max A. ; Pluim, Josien P.: Automatic nuclei segmentation in H&E stained breast cancer histopathology images. En: *PloS one* 8 (2013), Nr. 7, p. e70221

[54] Wang, Jun ; Liu, Qianying ; Xie, Haotian ; Yang, Zhaogang ; Zhou, Hefeng: Boosted efficientnet: detection of lymph node metastases in breast cancer using convolutional neural networks. En: *Cancers* 13 (2021), Nr. 4, p. 661

[55] Weaver, DL ; Krag, DN ; Manna, EA ; Gilcrease, Michael: Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. En: *Breast Diseases* 15 (2004), Nr. 2, p. 156

[56] Weaver, Donald L.: Sentinel lymph nodes and breast carcinoma: which micrometastases are clinically significant? En: *The American journal of surgical pathology* 27 (2003), Nr. 6, p. 842–845

[57] Weaver, Donald L. ; Krag, David N. ; Ashikaga, Takamaru ; Harlow, Seth P. ; O'Connell, Maureen: Pathologic analysis of sentinel and nonsentinel lymph nodes in breast carcinoma: a multicenter study. En: *Cancer: Interdisciplinary International Journal of the American Cancer Society* 88 (2000), Nr. 5, p. 1099–1107

[58] Weaver, Donald L. ; Krag, David N. ; Manna, Edward A. ; Ashikaga, Taka ; Harlow, Seth P. ; Bauer, Kenneth D.: Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. En: *Modern pathology* 16 (2003), Nr. 11, p. 1159–1163

[59] Wong, Sandra L. ; Faries, Mark B. ; Kennedy, Erin B. ; Agarwala, Sanjiv S. ; Akhurst, Timothy J. ; Ariyan, Charlotte ; Balch, Charles M. ; Berman, Barry S. ; Cochran, Alistair ; Delman, Keith A. [u. a.]: Sentinel lymph node biopsy and management of regional lymph nodes in melanoma: American Society of Clinical Oncology and Society of Surgical Oncology Clinical Practice Guideline Update. En: *Annals of surgical oncology* 25 (2018), Nr. 2, p. 356–377

[60] Wong, Stephanie M. ; Almana, Nora ; Choi, Jungeun ; Hu, Jiani ; Gagnon, Haley ; Natsuhara, Kelsey ; Shen, Abra H. ; DeSantis, Stephen ; Dominici, Laura ; Golshan, Mehra [u. a.]: Prognostic significance of residual axillary nodal micrometastases and isolated tumor cells after neoadjuvant chemotherapy for breast cancer. En: *Annals of surgical oncology* 26 (2019), Nr. 11, p. 3502–3509

[61] Yamamoto, Hirofumi: Micrometastasis in lymph nodes of colorectal cancer. En: *Annals of Gastroenterological Surgery* (2022)

[62] Zhang, Lu ; Hu, Jin ; Ali Zirakzadeh, A ; Rosvall, Jesper ; Hedlund, Mats ; Hu, Ping S. ; PA Wallin, Robert ; Sherif, Amir ; Winqvist, Ola: Detection of micro-metastases by flow cytometry in lymph nodes from patients with penile cancer. En: *BMC urology* 18 (2018), Nr. 1, p. 1–7

[63] Zhang, Yi ; Wang, Yuzhi ; Tian, Gang ; Jiang, Tianhua: Long non-coding RNA-based signatures to improve prognostic prediction of breast cancer. En: *Medicine* 99 (2020), Nr. 40