Aplicación de técnicas de preprocesamiento y segmentación de imágenes para el apoyo diagnóstico en la detección de Cáncer de Próstata

Julián David Vargas López

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería Electrónica
Manizales, Colombia
2022

Application of image preprocessing and segmentation techniques for diagnostic support in the detection of prostate cancer.

Julián David Vargas López

Thesis presented as partial requeriment to obtain the dedree:
Master of Engineering

Advisor:
Juan Bernardo Gómez Mendoza, Ph.D

Research Line:
Industrial Automation
Research Group :
Soft and Hard Applied Computing (SHAC)

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería Electrónica
Manizales, Colombia
2022

Dedicado a quien viajó mucho . . . ahora en el cielo. Gracias Tío Hernán

Acknowledgements

En primer lugar quiero agradecer a mi tutor, Juan Bernardo Gómez Mendoza, quien con sus conocimientos y apoyo me guió a través de este proceso formativo de la maestría; de igual manera quiero agradecer al profesor Andrŕes Marino Álvarez por los consejos brindados durante la maestría. También quiero agradecer a los doctores Alexander Pava Marín, Paula Andrea Toro y al profesor Nicolás Toro García por brindarme todo su apoyo y conocimiento que fueron necesarios para llevara a cabo el proceso de investigación. Por último, quiero agradecer a todos mis compañeros y a mi familia, por apoyarme aun cuando mis ánimos decaían. En especial, quiero agradecer a mis padres, que siempre estuvieron ahí para darme palabras de apoyo y un abrazo reconfortante para renovar energías.

Este trabajo de tesis fue apoyado por Minciencias mediante el proyecto de investigación Caracterización morfológica de estructuras cerebrales por técnicas de imagen para el tratamiento mediante implantación quirúrgica de neuroestimuladores en la enfermedad de Parkinson - Código: 11180763808; así como por la Universidad Nacional de Colombia y la Universidad de Caldas mediante el proyecto de investigación Construcción de una herramienta de apoyo diagnóstico en la detección de cáncer de próstata en imágenes – Código Hermes: 42271 .

Resumen

El aprendizaje profundo ha tenido un impacto notable en el análisis de imágenes médicas. Desde clasificar tejidos hasta localizar áreas anormales en una región, herramientas como las redes neuronales convolucionales (CNNs) y sus múltiples arquitecturas han mostrado resultados prometedores en esta área de la medicina. En patología digital, estos modelos neuronales se están convirtiendo cada vez más en una herramienta vital en el apoyo diagnóstico y pronóstico para los patólogos. Actualmente, múltiples instituciones médicas utilizan CNNs en sus laboratorios para optimizar el tiempo de búsqueda de regiones anormales en imágenes médicas digitales - como lo son las muestras de biopsias -, generando automáticamente información relevante en el diagnóstico y pronóstico de un paciente. Su aplicabilidad se ha logrado en gran medida gracias a la existencia de habilitadores tecnológicos, como hardware especializado (p.e., procesadores gráficos o GPUs), que permiten manipular y procesar grandes cantidades de datos de manera simultanea. Sin embargo, las GPUs no puede procesar las imágenes en algunos casos debido a su tamaño. Las imágenes histopatológicas son un ejemplo de este tipo de imágenes, donde el tamaño de las imágenes pueden ser del orden de hasta 25.000 x 30.000 píxeles. Se han diseñado estrategias que permiten manipular este tipo de imágenes, desde optimizar la forma de entrenar las CNNs hasta dividir la imagen en parches con un tamaño manejable. Sin embargo, analizar la biopsia, elegir las áreas de interés y crear las etiquetas correspondientes, son procesos que se realizan de forma manual y resultan dispendiosos para el especialista. Por lo tanto, es necesario desarrollar nuevas estrategias para apoyar al patólogo en estas tareas. En este documento, se plantean tres metodologías que permiten apoyar al patólogo en el análisis de imagenes histopatológicas de tejido prostático.

El primer diseño emplea transformaciones de color que proporcionan información adicional sobre la imagen. Se mostró cómo estas técnicas mejoran y resaltan las estructuras presentes en el tejido (se logra una mejor definición de los núcleos, aumenta el contraste en el estroma y las células epiteliales, etc.). Estas transformaciones de color tienen la ventaja de que su implementación no genera un costo computacional considerable, permitiendo manipular la imagen de forma rápida, incluso en ordenadores que no posean un hardware especializado. El segundo diseño analiza el proceso de segmentación de una imagen con redes neuronales convolucionales. Se expuso el problema que se genera cuando se trata de clasificar estas imágenes dividiéndola en pequeños parches, en donde el tiempo de segmentación por imagen puede llegar a las 24 horas o más. En consecuencia, se diseña una estrategia para mitigar este problema empleando un porcentaje de pixeles de la imagen para segmentarla. Esta técnica permite disminuir el tiempo de segmentación a solo 5 minutos por imagen. Además, se logró demostrar experimentalmente, que la información que se pierde a medida que se disminuye el porcentaje de pixeles es muy pequeña (cerca del 5%), en comparación con el proceso en donde se emplean todos los pixeles de la imagen. Finalmente, nuestro

tercer diseño consiste en crear una metodología que permite localizar las áreas sospechosas en imágenes de cáncer de próstata utilizando redes neuronales convolucionales. Empleando los resultados de la etapa anterior, se diseña una red neuronal convolucional que posee una cantidad pequeña de parámetros de entrenamiento (cerca de 50 mil). Esta red realiza dos tareas distintas: segmentar el estroma y segmentar el tejido sospechoso. Uniendo estos dos resultados y descartando los pixeles que pertenecieran al estroma segmentado, se logra localizar zonas sospechosas en imágenes de tejido prostático. Adicionalmente esta red se diseñó pensando en el costo computacional que generan algunas redes en el estado del arte, y en el sobredimensionamiento del problema que puede surgir al emplear dichas redes.

Palabras Claves: Tejido de próstata, Redes neuronales Convolucionales, optimización, Segmentación , Balance de Blancos

Abstract

Deep learning has had a noticeable impact on medical image analysis. From classifying tissues to locating abnormal areas in a region, CNNs and their multiple architectures have shown a future in this area of medicine. In digital pathology, these neural models are increasingly becoming a vital tool in diagnostic and prognostic support for pathologists. Currently, multiple medical institutions use CNNs in their laboratories to optimize the search time for abnormal regions of a complete tissue slide (biopsy sample), automatically generating diagnoses and prognoses of a patient, etc. This success of CNN was achieved mainly by using specialized hardware (GPUs) that allow large amounts of data to be manipulated and processed. However, analyzing the biopsy, choosing the areas of interest and creating the corresponding labels are processes that are carried out manually and are costly for the specialist. Therefore, it is necessary to develop new strategies to support the pathologist in these tasks. In this document, three methodologies are proposed that allow the pathologist to be supported in the analysis of histopathological images of prostate tissue.

The first layout employs color transformations that provide additional information about the image. It was shown how these techniques improve and highlight the structures present in the tissue (the shape of the nuclei was better defined, the contrast increased in the stroma and epithelial cells, etc.). These color transformations have the advantage that their implementation does not generate a considerable computational cost, allowing the image to be manipulated quickly, even on computers that do not have specialized hardware. The second design analyzes the segmentation process of an image with convolutional neural networks. The problem generated when we try to classify these images by dividing them into small patches, where the segmentation time per image can reach 24 hours or more, was exposed. Consequently, we designed a strategy to mitigate this problem by using a percentage of pixels in the image to segment it. This technique allowed the segmentation time to be reduced to only 5 minutes per image. In addition, we were able to demonstrate experimentally that the information lost as we decrease the percentage of pixels is very small (about 5%), compared to the process where all the pixels of the image are used. Finally, our third design creates a methodology that locates suspicious areas in prostate cancer images using convolutional neural networks. Using the previous stage results, we design a convolutional neural network with a small number of training parameters (about 50 thousand). This network performs two distinct tasks: segmenting the stroma and suspect tissue. Combining these two results and discarding the pixels that belonged to the segmented stroma, it is possible to locate suspicious areas in images of prostate tissue. Additionally, this network was designed considering the computational cost generated by some networks in the state of the art and the over-sizing of the problem that can arise when using these networks.

Keywords: Prostate tissue, Convolutional Neural Networks, Optimization, Segmentation, White Balance technique

Contents

Acknowledgements **vii**

Abstract **ix**

1 Introduction **xv**

2 Color enhancement of Histological Prostate Image **xix**
2.1 Image Enhancement xix
2.2 White Balance xx
2.3 Color transformations xxii
2.4 Algorithm evaluation methodology xxiv
2.5 Results xxv
2.6 Conclusion xxviii

3 Incremental Image Segmentation Using CNNs **xxix**
3.1 Introduction xxx
3.1.1 Basic architecture of a CNN xxx
3.2 Inference problem in images with CNN. xxxii
3.3 Proposed method xxxiv
3.3.1 Database xxxiv
3.3.2 Preprocessing xxxv
3.3.3 Network Architecture xxxvi
3.3.4 CNN Training and Validation xxxvii
3.3.5 Inference - Proposed Algorithm xxxix
3.4 Results xl
3.5 Conclusion xlv

4 Location of suspicious areas using CNNs **xlvi**
4.1 Introduction xlvi
4.2 Proposed methodology l
4.2.1 DataBase l
4.2.2 Preprocessing li
4.2.3 Architecture Network li

4.2.4 Stroma segmentation li
4.2.5 Location of abnormal areas lii
4.3 Results liv
4.3.1 Stroma Segmentation liv
4.3.2 Location of abnormal areas lvii
4.4 Conclusion lvii

5 Conclusions **lix**

6 Future work **lx**

Bibliography **lxi**

1 Introduction

In Colombia, prostate cancer occupies the first place in cancer incidence in the male population. In 2017, the mortality rate due to prostate cancer was 14.9 deaths per 100,000 men[1]. Prostate cancer screening tests can help find cancer early, making it easier to treat. A frequent form of cancer detection is analyzing the PSA (Prostate-Specific Antigen) found in the patient's blood. Another form of detection is through digital rectal examination. When the results of these tests indicate some anomaly, a small sample of prostate tissue (biopsy) is extracted to be analyzed under a microscope [1].

Pure samples cannot be analyzed directly as they have a transparent coloration typical of the tissue. Therefore, for proper analysis, pathologists use stains to color the tissue. One of the most used stains in histological tissues is the H & E stain [2]. This staining allows cell nuclei to be dyed in dark purple; the stroma or muscle tissue becomes pinkish, the red blood cells turn light red, and the cytoplasm becomes light purple. An example of this staining can be seen in Figure **1-1**.

To determine the presence of cancer in the tissue sample, pathologists must examine each area of the image looking for the morphological characteristics of the glands present in the sample. In the 1960s, a scale was established to indicate the degree of aggressiveness of prostate cancer. This indicator is known as the Gleason Scale. This is based on the normal architecture of the glandular tissue (shape, size and differentiation of the glands)[3]. Figure **1-2** shows the Gleason scale. Diagnosis procedure based on the Gleason scale consists of dividing an area of the image into two portions. Each zone is given a value between 1 and 5, where a value of 1 corresponds to mildly aggressive or benign tumors and 5 to tumors with a high degree of cancer. Once the two values have been assigned to the area of interest, these values are added. In Table **1-1** we can see the interpretation for each of the sum ranges.

Thanks to the emergence of digital scanners, the histological sample can be digitized to produce a high-resolution image. This allows for better visualization of tissue, faster scrolling through the sample, and brought ease to sharing information between pathologists in search for a second opinion. However, to give a complete patient diagnosis, the pathologist must

[1]Instituto Nacional de Cancerología ESE. Análisis de la Situación del Cáncer en Colombia 2015. Primera edición. Bogotá D.C: Instituto Nacional de Cancerología ESE

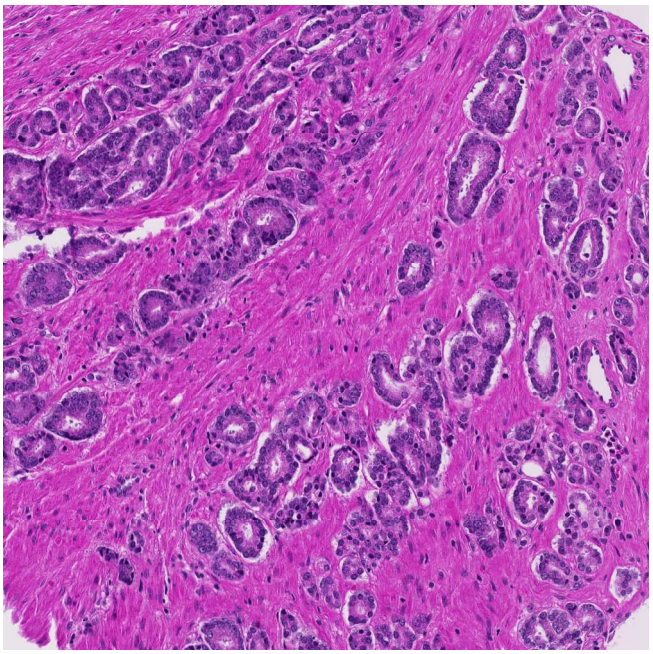


Figure **1-1**: Example of an image with H & E staining. Pink color corresponds to muscular tissue, purple to epithelium, and dark purple to cells.

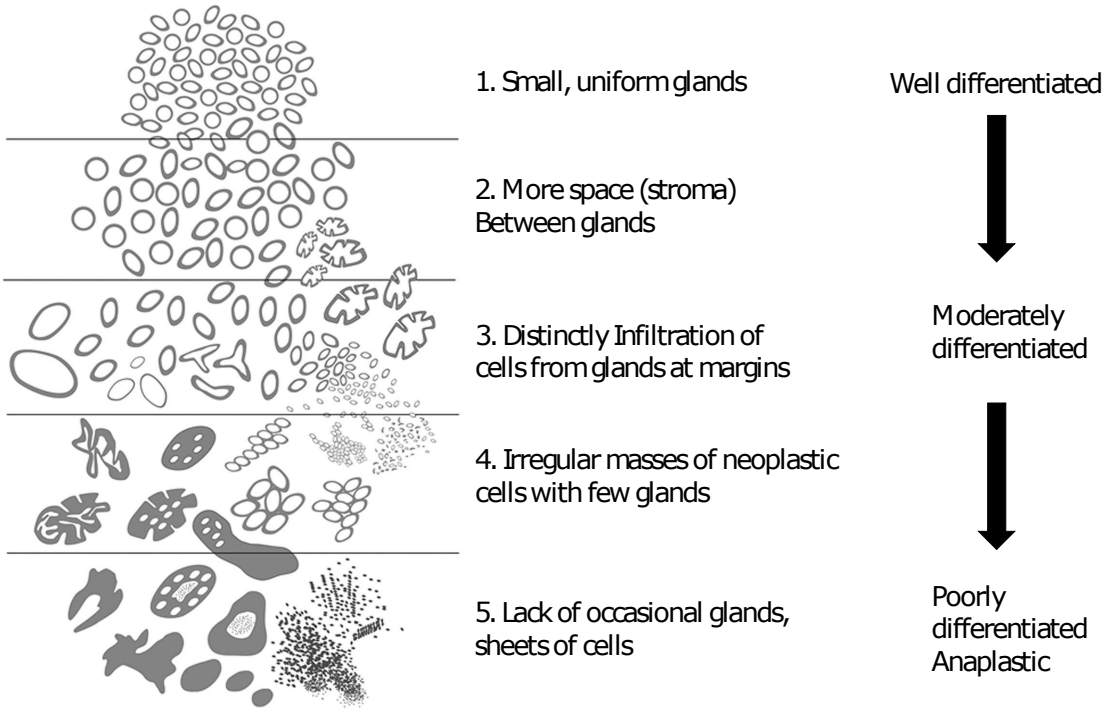


Figure **1-2**: Gleason scale - adapted from [4].

Gleason Score	
Gleason Scale	**Interpretation**
Gleason 2 - 6	Well-differentiated tumor tissue, little presence of cancer, better prognosis.
Gleason 7	Moderately differentiated tissue, gland infiltrates
Gleason 8 -10	Poorly distinguishable tumor tissue, aggressive cancer, worse prognosis

Table **1-1**: Gleason pattern

analyze the sample completely. This process is complex and time-consuming [5], so it is necessary to use new methodologies to reduce the time spent in this type of analysis. In the last decade, new computational techniques have been designed to analyze digitized samples using Computer Vision techniques.

For instance, in [6, 7, 8, 9, 10, 11], the authors designed different methodologies based on Machine Learning techniques that allowed establishing differences between the morphology of the tissues in the images and the Gleason patterns automatically. They used low-level characteristics such as the proportion of the gland perimeter, the area occupied by a gland, the distances from the center of the gland to the border, etc. The authors used Support Vector Machines (SVMs) to classify tissues, particularly because it presented advantages such as having a higher performance index (precision and accuracy) with a good level of robustness, and allowing faster results than CNNs.

In some cases, it is not only required to differentiate the Gleason scales in an image, but it is also necessary to detect the gland and classify it. Hua et al. [12] introduces a methodology in which co-occurrence matrices and color histograms are used as the model's main features. The segmentation is performed with SVMs, achieving a gland detection rate of 86%.

Mahoor et al. [13] proposes the use of a Shearlet filter bank to detect whether a gland is benign or malignant. Like the Gabor filter bank, Shearlet filters allow you to extract edge and texture features using different rotations. The classifier used was the SVM, with an accuracy rate of 86%. Tabesh et al. [14] used texture characteristics, color, and morphological descriptions to classify the tissue into areas with a high presence of cancer and regions with a low presence of cancer. The model was trained with the k-nearest neighbors, Gaussian and SVM classifiers, the latter with the best detection rate: 81%. Kumar et al. [15] designed a classifier with 4 features: color channel histograms, fractal dimension features, fractal code features, and wavelet features. The purpose was to detect the presence of tumors in the tissue (tumor class and non-tumor class). The model was trained with Gaussian linear and quadratic classifiers, where the quadratic classifier obtained the best performance: 77% tumor detection.

In the previous works, feature extraction is a determining factor in the model's performance.

However, it is necessary to invest time in identifying the characteristics that allow to model an image, which can generate subjectivity problems. Since the advent of CNN, not only has this problem been eliminated, but ranking performance has increased. In [16, 17, 18, 19, 20, 21], the authors present different methodologies for gland segmentation based on CNNs. The purpose is to detect, segment, and classify whether or not the gland in question is cancerous. The indices obtained when carrying out this process were higher than 92%, surpassing the techniques mentioned above.

Although CNN models are robust, some tasks have a high computational cost ((for example, when it comes to going through an image and detecting or classifying the pixels), and the time required for both training and validation can take hours. In this document, we aim to present three tools that allow the pathologist to be supported in the task of making the diagnosis. Chapter 1 shows a series of color transformations that enable the tissue to be visualized with better detail than the original image. In chapter 2, we expose a bottleneck when CNNs are used to segment images by patches, in which we design a novel solution (a different perspective in the way of going through an image) that considerably mitigates this problem. And finally, in chapter 3, we use the results of the previous chapter to create a methodology to locate regions with a high degree of cancer in prostate images. Finally, in the future work section, we propose some pending tasks in the framework of this research.

2 Color enhancement of Histological Prostate Image

Tissue imaging classification is one of the most important steps in prostate cancer diagnosis. The pathologist relies on the gland's structure to assign a Gleason grade to the tissue image. However, due to the limitations inherent to the acquisition process, the images have low illumination and low contrast, which significantly influences the analysis of the tissue. Therefore, it is required to transform these low-quality images to highlight features that ease the pathologist's diagnostic process.

This Chapter presents some image processing techniques focused on enhancing contrast in histopathological images of prostate tissue. The proposed approach consists of the implementation of 44 classical color transformations without the application of any filter. A group made up of 3 pathologists will analyze each of these transformations to select those that provide relevant information for the diagnosis of prostate cancer. This Chapter is divided into the following sections. Section 2.1 describes contrast manipulation, histogram manipulation, and color manipulation focused on enhancing image contrast. Section 2.2 describes the white balance technique, which is one of the most widely implemented algorithms in histopathological imaging. Section 2.3 enunciates the 44 color transformations implemented in this work, and their application in state-of-the-art. Section 2.4 addresses the methodology implemented for the group of 3 pathologists to analyze each of the color transformations shown in Section 2.3. In Section 2.5 we present the results of the analysis and the transformations selected by the pathologists, and finally, in Section 2.6 we present the conclusion of the Chapter.

2.1 Image Enhancement

Image enhancement processes consist of a collection of techniques that seek to improve the visual aspect of an image or convert it to a more suitable form for analysis by a human or a machine [22]. The digital image enhancement techniques provide a multitude of options to improve the visual quality. The imaging modality greatly influences the appropriate choice of

such methods, the task in question, and the visualization conditions [23]. For the realization of this Thesis, three classes of algorithms were used [24]:

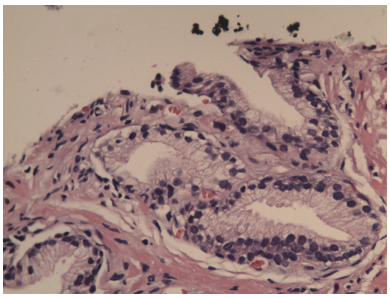
- **Histogram manipulation:** The histogram is a graph that indicates how often each degree of light intensity or color in the image is displayed. The intensity levels are generally represented by numbers ranging from 0 to 255, with 0 denoting a low intensity (the lowest) and 255 denoting a high intensity (the highest). This approach involves altering the values in an image's histogram to boost the image's contrast. The algorithms selected from this section were: the histogram equalization and the adaptive equalization of the histogram.
- **Contrast manipulation:** One of the most common image defects is low contrast, which results from a narrow and possibly nonlinear image width range. Some objects present in the image cannot be easily or correctly identified. This technique consists of the direct manipulation of the pixels of an image with linear and nonlinear functions that make it possible to readjust the intensity levels of the image. Implementing these algorithms is observing the objects with better clarity than before, and they could be due to the lighting problem not being detailed. The algorithm chosen for the document is white balance.
- **Color Manipulation:** Also known in the literature as color transformations. These techniques make it possible to change the initial color space to a different color space to observe in whole or in part (through the newly created bands) features that emphasize the content of the original image [25]. Assuming the original image in RGB representation, the transformations selected in this section were: XYZ, YCrCb, I1I2I3, CMYK, YIQ, Lab, HSI, and HMMD [26].

The justification for selecting these transformations can be seen in the Section 2.3

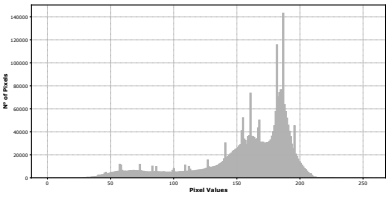
2.2 White Balance

In digital image processing, white balance is one of the most popular contrast manipulation techniques. It involves applying a linear transformation to the image in order to balance the color predominance and alter the image intensity. The histogram reflects this modification, with intensity values stretched to extremes, allowing the image's contrast to dramatically improve. Figure **2-1** shows a histological picture of benign prostate tissue with histograms (one for each channel). In Figure**2-1**a, the image shows a low contrast that is reflected in the histogram of each RGB channel. In Figures **2-1**b, **2-1**c and **2-1**d we find that the pixels are grouped in a range that does not cover the extremes; that is, none of the pixels is white (value of 255) or black (value of 0). Figure **2-2**a shows the image with the white balance

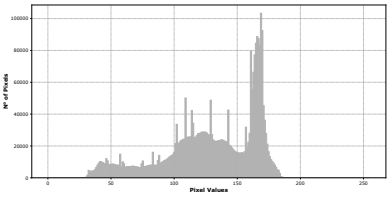
algorithm. Figures **2-2**b, **2-2**c and **2-2**d show the stretching of the histogram. Notice that white balance does not change the overall shape of the histogram.



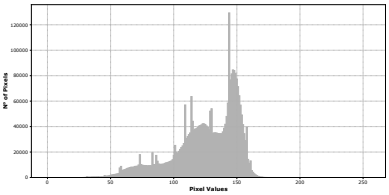
(a) Original image.



(b) Red channel histogram.

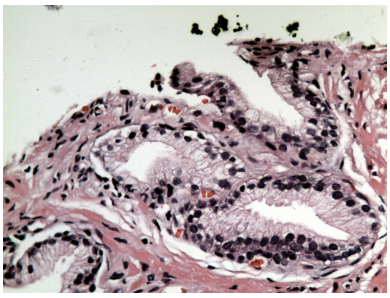


(c) Green channel histogram.

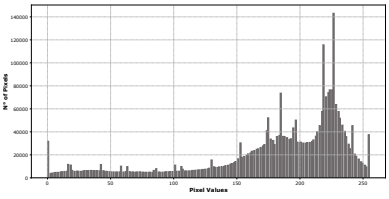


(d) Blue channel histogram.

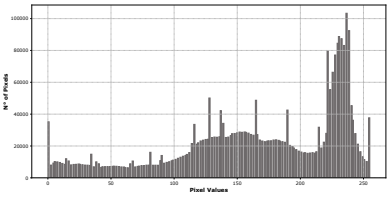
Figure **2-1**: Image of prostate tissue categorized as benign, along with its respective RGB channel decomposition.



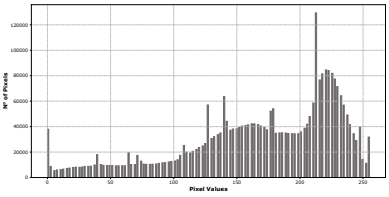
(a) Balanced image.



(b) Red channel histogram.



(c) Green channel histogram.



(d) Blue channel histogram.

Figure **2-2**: Image of prostate tissue after applying the white balance algorithm, in addition to its respective RGB decomposition.

According to [27], white balance can be achieved by computing:

$$C_{wb} = L_{out} + (H_{out} - L_{out}) \left(\frac{I_{x,y} - L_{in}}{H_{in} - L_{in}} \right)^{\gamma} \tag{2-1}$$

where:
C_{wb} is the image that results after applying white balance.
L_{in} is the smallest pixel value that in this case is given by the 1st percentile of the image.
L_{out} is the value of the smallest output pixel you want to reach, which in this case is 0.
H_{in} is the largest pixel value that is given by the 99th percentile of the image.
H_{out} is the value of the largest output pixel you want to reach, which in this case is 1.
$I_{x,y}$ is the original low contrast image.
γ is the gamma correction factor that in the white balance technique is equal to 1.

2.3 Color transformations

In works such as [28, 29, 30] employ different types of color models that are used as characteristics or parameters that facilitate the segmentation or classification of prostate tissue. The use of these color models is because they present some relevant information regarding the area to be classified. Given the results obtained in these works, it was proposed to use some color spaces, together with histogram and contrast manipulation techniques, which would allow the pathologist to reveal visual information regarding the original image extracted from the microscope. Considering the amount of color manipulation, histogram, and contrast techniques that exist, it was decided to choose those that were easy to implement or were previously implemented in some programming language, in addition to having been used in other scientific articles. The list of algorithms can be seen in Table **2-1**. The 44 techniques were divided into three new categories that describe the nature of each one to analyze each one of them:

- Color Transformations using RGB
- Linear Transformations based on RGB
- Non-linear Color Transformations.

The techniques shown in Table **2-1** were selected thanks to different works framed in the area of digital image processing detailed below.

Baldevbhai et al. [31], implemented the XYZ and L * a * b color spaces for the segmentation of medical images under the clustering paradigm. The authors used only the pixel intensity characteristics of the image with which they designed a new methodology for segmentation.

Gasparri et al. [32], used the HSI color space as a base technique to develop a new method of automatic segmentation of vasculature in retinal images for the study of diabetic retinopathy. The authors recommend the use of the HSI color space since this space can separate the

Linear Transformations based on RGB	**Non-linear Color Transformations**	**Color Transformations using RGB**
XYZ + WB	H → HSI	Histogram Equalization (E)
Y → YCbCr	S → HSI	Adaptive Histogram Equalization (AE)
Cb → YCbCr	I + WB → HSI	White Balance (WB)
Cr → YCbCr	H → HMMD	Channel R → RGB
I1 + WB → Hearing	Min + WB → HMMD	Channel G → RGB
I2 + WB → Hearing	Max + WB → HMMD	Channel B → RGB
I3 → Hearing	D + WB → HMMD	Channel R + WB
C + WB → CMYK	L + WB → Lab	Channel G + WB
M + WB → CMYK	a → Lab	Channel B + WB
Y → CMYK	b + WB + AE → Lab	Channel R + E
K → CMYK	u → Luv	Channel G + E
Y + WB → YIQ	v → Luv	Channel B + E
I + WB → YIQ		Channel R + AE
Q → YIQ		Channel G + AE
U → YUV		Channel B + AE
V → YUV+WB		WB + AE

Table **2-1**: List of color transformations used.

intensity of the essential information from the color even when they are affected by the lighting changes, a problem they had with the images used.

Wurood et al. [33] applied histogram equalization and adaptive histogram equalization techniques to improve the contrast of medical magnetic resonance imaging (MRI). The authors reported excellent efficiency to improve the contrast of MRI images with high quality. For this, they considered quality measures such as MSE, PSNR, and SNR to evaluate the quality of the improved medical MRI images.

Singh et al. [34], created a new technique to improve image contrast using discrete wavelet transform and Luv color space. The result of this new technique is better quality images with less noise and blur.

Lastly, the HMMD, YCbCr, I1I2I3, CMYK, YIQ and YUV color spaces were used by Arash Abadpour [26], where he made the first experimental evaluations to compare the representation based on the analysis of main components (PCA) of the images in color of nature with the respective representations based on the color space. This framework is used to design color image processing algorithms with superior performance, such as coloring, compression, segmentation, deliberate distortion, etc.

2.4 Algorithm evaluation methodology

The rating scales consist of a series of qualitative categories in which an observer must issue a judgment, indicating the presence of a characteristic. One of the most used rating scales is the Likert Scale [35]. Although many of the evaluations are subjective, they determine some patterns or behaviors according to the subject being evaluated.

An image quality test was used to analyze each of the 44 color transformations. The procedure consisted of showing four images chosen at random from the set of images to 3 pathologists, of which two have experience in the area of urogenital pathology. A color transformation was applied to each of these images (the same transformation for the 4).Each pathologist rated the images on a scale of 0 to 4, where:

- **Qualification 0:** The transformation does not provide any information to the pathologist.
- **Qualification 1:** The transformation provides little information, but even so, it is discarded by pathologists.
- **Qualification 2:** The transformation provides little information, but does not high-

light, in some cases, the areas of interest for the pathologist.

- **Qualification 3:** The transformation provides valuable information, but, in some cases, fails to highlight areas of interest to pathologists.
- **Qualification 4:** The transformation provides information of great value to the health specialist. Areas of interest for pathologists are highlighted.

2.5 Results

In the process of qualifying color transformations, pathologists chose 23 techniques, that is, 52.27% of our algorithm list. These obtained a score of 3 or 4 on our scale, qualifying them as useful for a better diagnosis. The percentage of highlighted techniques was calculated by dividing the transformations that had an excellent performance, that is, those that had a rating of 3 or 4, by the total number of transformations that were used.

The results of the qualification carried out by the pathologists for each group of algorithms (linear transformations based on the RGB space, non-linear color transformations and transformations using the RGB space) can be seen in Figures **2-3**, **2-4** and **2-5**:

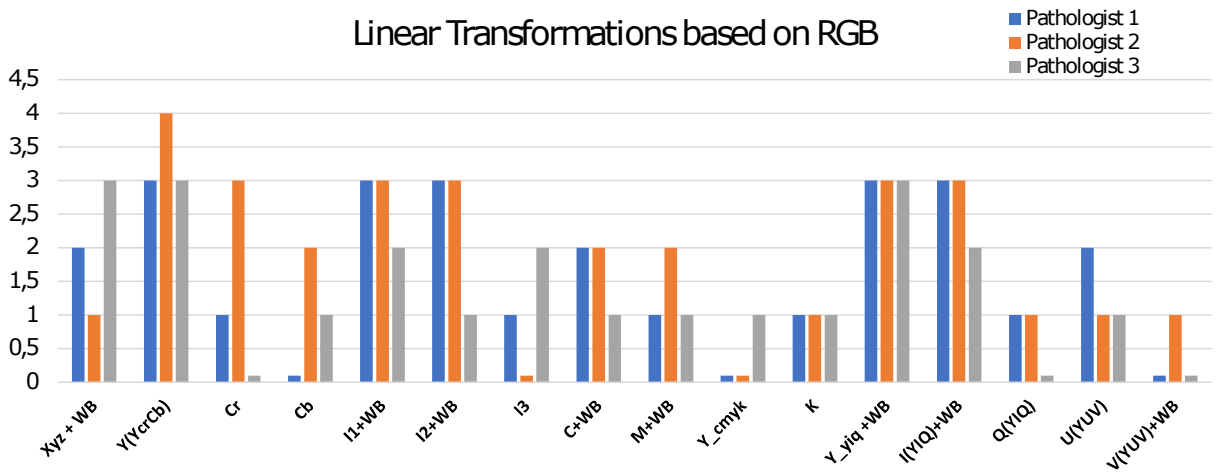


Figure **2-3**: Graphs of the qualifications by the three pathologists corresponding to the linear transformations based on the RGB color space.

Table **2-2** shows the color transformations that pathologists classified as useful. These allowed highlighting the structures present in the images of prostate tissue such as cytoplasm, stroma, and nuclei. The white balance, which was the best-qualified transformation, managed to improve the perception of the contour and the shape of the nuclei.

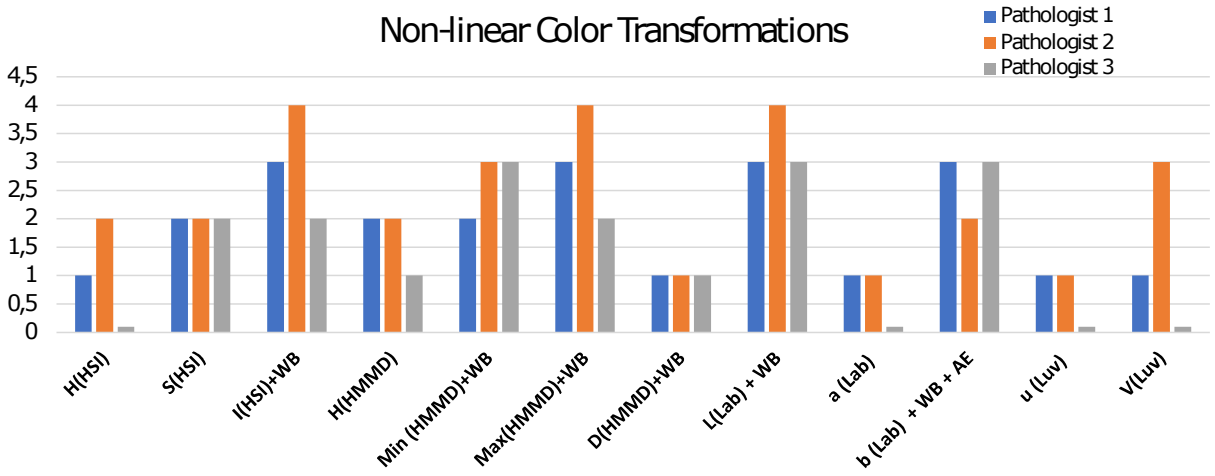


Figure **2-4**: Graphs of the qualifications by the three pathologists corresponding to the non-linear transformations.

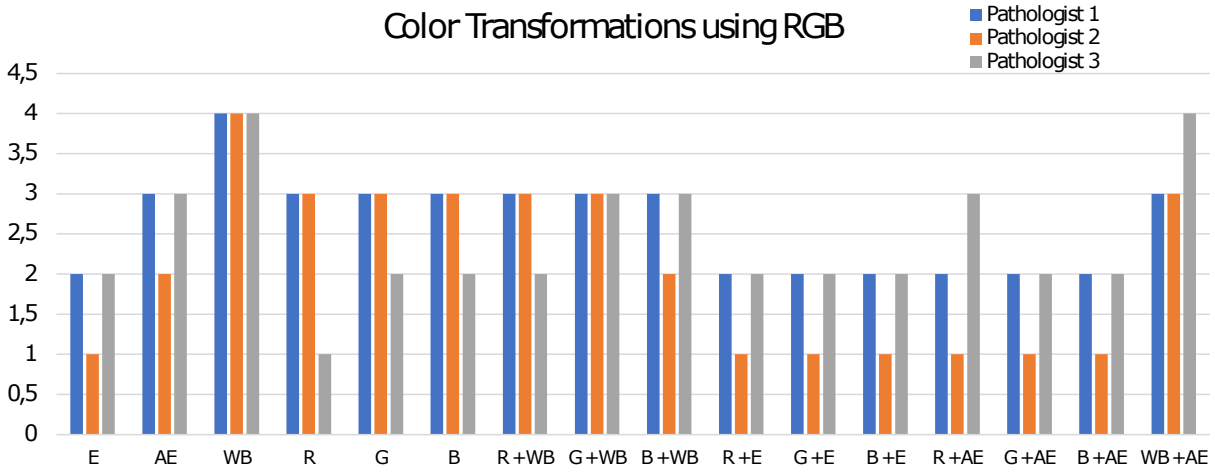
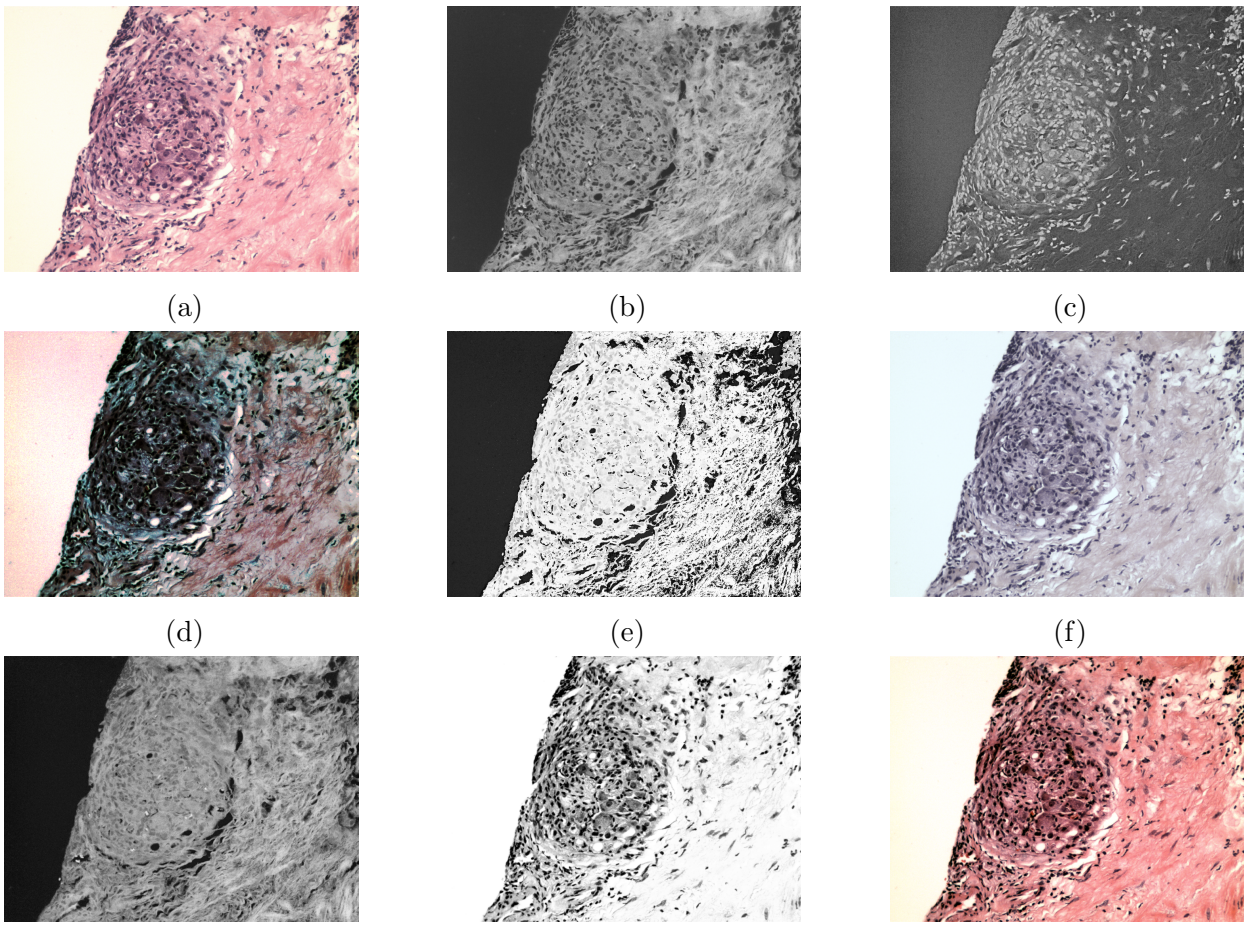


Figure **2-5**: Graphs of the qualifications by the three pathologists corresponding to the linear transformations using the RGB color space.

Linear Transformations based on RGB	**Non-linear Color Transformations**	**Color Transformations using RGB**
Y(YIQ) + WB	L(L*a*b) + WB	WB + EA
I(YIQ) + WB	b(L*a*b) + WB + EA	R(RGB) + EA
I1 + WB	V(Luv)	R(RGB) + WB
I2 + WB	Min(HMMD) + WB	G(RGB) + WB
Y(YCrCb)	Max(HMMD) + WB	B(RGB) + WB
Cr(YCrCb)	I(HSI) + WB	R(RGB)
XYZ + WB		G(RGB)
		B(RGB)
		WB
		EA

Table **2-2**: Transformations summary



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure **2-6**: Example of rating an image

Figure **2-6** shows different images qualified by pathologists. Figure **2-6**a corresponds to the original image of a sample of prostate tissue of a Gleason 5 pattern with a magnification of 20x.

Figures **2-6**b and **2-6**c show color spaces whose rating was 0. In these, we can see respectively the V (YUV) color space plus white balance and Cb (YCbCr) color space. For pathologists, these types of transformations did not provide any information relevant to the image enhancement process. On the contrary, they deleted information contained in the original image.

Figures **2-6**d, **2-6**e, **2-6**f and **2-6**g are color spaces whose ratings varied between 1 and 2. In these, we can see the image equalization, H (HIV) color space, XYZ color space and a (L*a*b) color space, respectively. Although some of these transformations do not provide all the complete information related to the diagnosis, the prognosis, and the predictive factors if they highlight the cell nuclei and allow us to see the basal cells that are important for the diagnosis of pathology.

Figure **2-6**h is the color space whose ratings varied between 3 and 4. In these image, we can see Red (RGB) color space plus white balance. These color space allowed the pathologists to visualize with better clarity the nuclei present in the tissue. Also, it show important characteristics for the diagnosis of pathology of prostate tissue managing to provide more benefits than the microscope.

Finally, Figure **2-6**i is the transformation that coincided in the evaluation with four by the pathologists and corresponded to the white balance. This technique allowed them to observe with a better contrast all the structures present in the images, adding important information to perform the corresponding diagnosis.

2.6 Conclusion

The color transformations used in this work not only improved the visual quality of the image but also highlighted some areas of great importance for the diagnosis of prostate cancer. For example, the nuclei were highlighted in a way that the perception of the contour and shape was improved. Linear color transformations and transformations based on RGB space were the ones that scored the best. The technique that received a rating of 4 unanimously from pathologists was the white balance, as it accentuated the image structures much better with regard to the other transformations. Finally, the methodology used in this section managed to cover a fundamental problem in the area of images contrast of the prostate tissue to help in the diagnosis and evaluation of this disease.

3 Incremental Image Segmentation Using CNNs

In the previous chapter, we showed some color transformations to improve the contrast of structures such as nuclei, cytoplasm, and stroma. These color spaces provided a visual tool for pathologists to easily identify areas of interest and rule out tissues that are not considered malignant. In recent years, many diagnostic support systems have processing techniques to improve tissue visualization and use machine learning algorithms to classify, detect, and segment histological images. For example, this is the case with convolutional neural networks. However, including these new methodologies implies new challenges in terms of computational cost to segment histological images since they are of considerable size.

In this chapter, we explain one of the challenges involved in using CNNs in histopathological images and how it can be solved by designing a new segmentation methodology. We proposed an algorithm that optimizes the segmentation time utilizing a percentage of image pixels. Additionally, we show that as we decrease the rate of pixels used to segment the image, the network's performance does not vary much compared to the use of all the image pixels. This chapter is divided into four sections. Section 1 introduces the main components of a convolutional neural network: the convolutional layer, the pooling layer, the perceptual field, and the fully connected layer. Section 2 exposes the classification problem in histological images using CNN. In addition, we show an alternative to mitigate the classification problem where we propose the new methodology to segment areas of interest in prostate cancer images. Section 3 presents the results of this technique, where we expose the impacts of the segmentation with different percentages of pixels and experimentally demonstrate that the method does not produce significant errors with the different rates of pixels. Finally, section 4 presents the respective conclusions of the chapter.

3.1 Introduction

CNNs use image recognition and classification to detect objects, recognize faces, identify abnormalities in medical images, and more [36]. CNNs are primarily used to classify images, group them by similarities, and then recognize the objects present in them. They are made up of neurons with weights and biases that can be learned. Each specific neuron receives numerous inputs and then computes a weighted sum that is passed through an activation function [37].

A CNN has two stages:

1. Feature extraction
2. Classification

The first stage consists of the automatic feature extraction, in the form of convolution masks [38]. This was a breakthrough in feature engineering. In classical Machine Learning techniques (such as SVM, decision trees, etc.) [39], the user was required to select the characteristics he considered relevant a priori. This generated two important drawbacks: determining the number of characteristics necessary to model a classification problem (this implied that the user had to know very well about the topic being investigated or hire expert personnel on the subject), and the introduction of bias or human error. Thanks to the new methodology of extracting characteristics utilizing convolutions, these two disadvantages could be considerably mitigated. The second stage uses a multilayer perceptron that uses the information extracted in the first stage to solve the classification problem.

3.1.1 Basic architecture of a CNN

Convolutional neural networks have three distinct layers: input, hidden, and output layers. In the input layer, we define the perceptual field that the neural network must use (the size of the input data). The hidden layers are where the convolutions (feature extraction) are performed. In the output layer is where we find the final CNN ranking. Next, we will detail each of the contents of each layer.

Perceptual Field (Input Layer)

In general, we define the perceptual field as the size of the window that will allow us to extract the features in the convolutional layer. In order to initialize the convolutional neural

network, we must first establish the size of the input data in tensor form (image height x image width x image channels). Once the size of the perceptual field is defined, it cannot be modified during the training and validation of the network. For example, if a perceptual field of 32x32x3 pixels is defined, the images to validate the convolutional model must be 32x32x3 pixels in size. The selection of the perceptual field size is made depending on the problem to be solved, and the capacity of the underlying hardware.

Convolutional Layer (Hidden Layer)

The convolutional layer is responsible for extracting the features of the input layer or the layers prior to it. Usually, a dot product of the convolution kernel is performed with the input matrix of the layer, as shown in Equation 3-1.

$$G[m,n] = (f * h)[m,n] = \sum_j \sum_k h[j,k] f[m-j, n-k] \tag{3-1}$$

Where $G[m,n]$ is the feature map; f represents the image or the input matrix; h is the convolution kernel; m, n are the row and column of the image, respectively; and j, k are the accumulation indexes.

As the convolution kernel slides along the input matrix, the operation generates a feature map (Figure **3-1**), which becomes the input of the next layer.

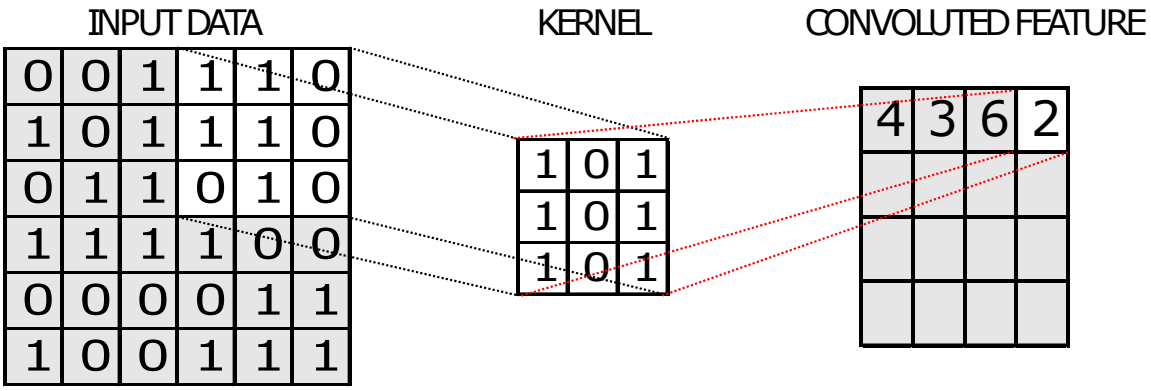


Figure **3-1**: Example of the convolution operation.

Pooling Layer (Hidden Layer)

Pooling layers allow us to reduce the dimensions of the feature map by grouping a set of neurons into a single one (local pooling) and positioning it in a new feature map. There are

two common types of pooling: maximum and average. Maximum pooling uses the maximum value of each set of neurons in the feature map, while average pooling uses the average set of neurons. In Figure **3-2**, we can see in detail the application of the pooling layer in the feature map.

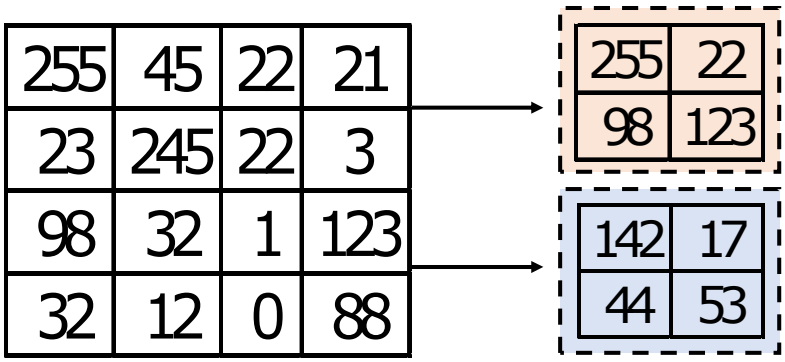


Figure **3-2**: Example of the pooling layer over a matrix. The red area corresponds to max-pooling operation, and the blue area corresponds to average-pooling operation.

Fully Connected Layer (Output Layer)

This layer uses the final feature map as input to a neural network known as a multilayer perceptron [40]. This network is used as a classifier, where the last set of neurons is equivalent to the set of classes to be classified in terms of probabilities. In figure **3-3**, we can see an example of fully connected layer architecture.

3.2 Inference problem in images with CNN.

With architectures like those in Section 1, CNNs can solve many classification problems with a high percentage of accuracy. For example, in imaging, convolutional neural networks were used to determine if a patient who suffered from COVID-19 has any condition in the lungs [41]. To do this, they use computed tomography images of the chest for the training and validation of the neural model. Another example in this area is using CNNs to detect pneumonia in X-ray images (around 5,800 images) [42]. Many similar articles in the literature use CNN to classify medical images.

However, in large images with a lot of information such as histopathological images, this form of classification is not possible. Normally, in the previous examples, images smaller

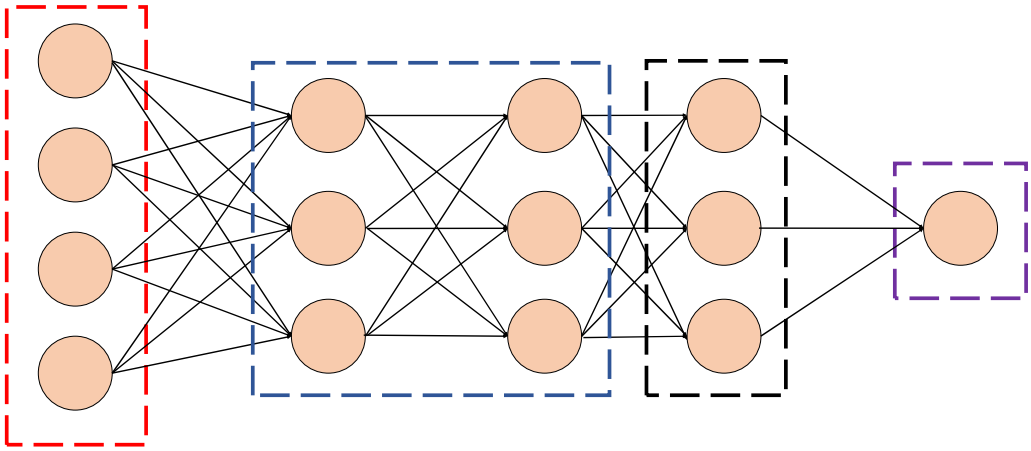


Figure **3-3**: Example of a fully connected network. The red box corresponds to the input of the network (it consists of 4 neurons); the blue box corresponds to the hidden layers of the network (it consists of two intermediate layers each of three neurons); the black box corresponds to the output of the network (consists of 3 neurons) and finally the purple box corresponds to the loss function of the network.

than 512x512 pixels are used, that is, they use the entire image to classify. In histological images, we have much larger images, with 3 channels each. Suppose we want to classify images of prostate tissue into benign or malignant tissue. We must design a deep network so that it can extract enough information which generates a huge number of parameters (around 1×10^{30} or more) that, even having the most powerful GPU, it will not even be able to train the network.

Additionally, for a CNN to achieve good accuracy rates, a representative data set [43] is required. In the histopathology databases, we can find a data set of up to 900 images with their respective annotations, with which, using a CNN, we will not be able to obtain a good performance in the classification.

Not all computer vision problems are posed in terms of classification (categorizing a set of data). It is often required to locate or segment a particular area in the image. In the analysis of histological images, this task is the most frequent. From this perspective, it is not necessary to work with a complete image, it is enough to divide the image into small patches, then select the patches that belong to a particular class and finally train the neural network with these small images.

This is a huge advantage. First, we eliminate the problem of missing data since by dividing an image into small patches, we increase the size of the training and testing data set, and

improve the performance of the neural network. Second, we attenuate considerably the large perceptual field problem since the perceptual field is reduced to the size of the patch, thus reducing model size.

However, this new form of segmentation creates a new bottleneck. Once the training and validation of the neural model is finished, the next step is to verify the results on an image. To do this, the image to be segmented is traversed using a window of the same size as the patches [44]. Each window is evaluated by the network, thus obtaining a label for each window. At the end of this operation, the segmented image will be obtained. Training a CNN for segmentation requires far fewer images than image recognition since any real-field image with the corresponding labeling can be used to obtain multiple training patches, and the CNN itself has fewer parameters to tune. Although this scheme uses a much smaller CNN than those used in image recognition, evaluating all the resulting sliding windows (nearly the same number of pixels in the image to be segmented) becomes much more demanding than evaluating a larger model using the entire image at once.

In the next section we will show the solution that was designed to solve the bottleneck by solving a problem of segmentation of the muscular structure (stroma) in images of prostate tissue.

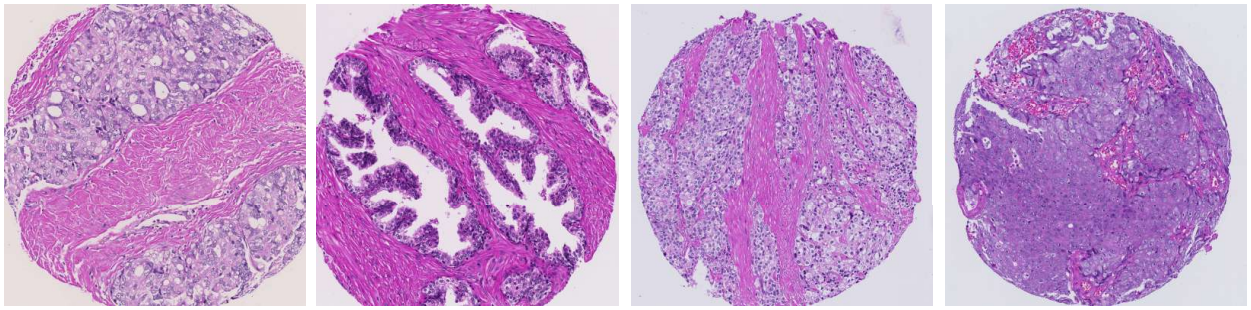
3.3 Proposed method

As mentioned above, to solve this problem, the following task was proposed: segment the stroma from the rest of the structures present in images of prostate tissue. The idea arose from the work of [45], where they used the super-pixel technique to segment the epithelium and stroma in histopathological images (tissue microarrays) of oropharyngeal tissue. Our approach is to separate muscle tissue using convolutional neural networks with our database.

3.3.1 Database

The dataset consists of 100 images in microarray format (See Figure **3-5**b) of prostate tissue, which were acquired at 40x optical magnification (0.23 microns per pixel) with a digital camera (NanoZoomerXR digital slide scanner) producing 3100x3100 pixel images of size. Our database is a subset of [46] consisting of 641 images in a microarray format. Our data subset consists of 25 benign tissue images, 25 Gleason 3 images, 25 Gleason 4 images, and 25 Gleason 5 images. The original database labels were not used (the authors' job was to detect suspicious areas and classify them in one of the four existing categories), but the

(a) Label images



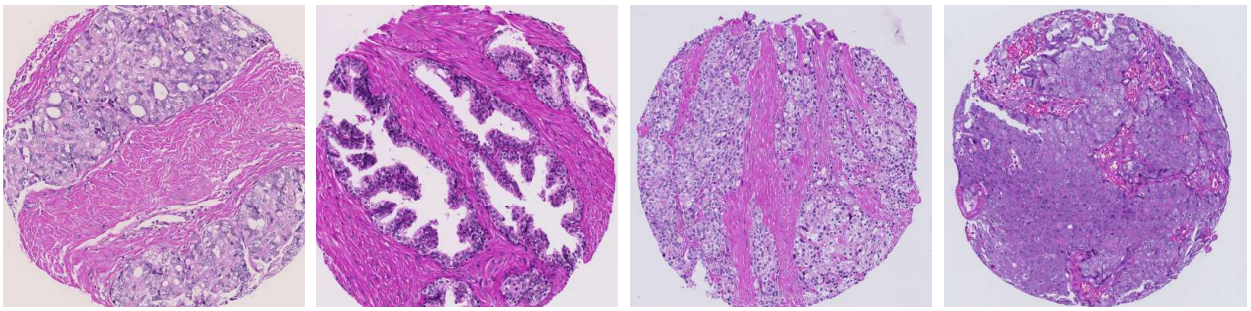
(b) Original Images

Figure **3-4**: Example of our database with its labels and images

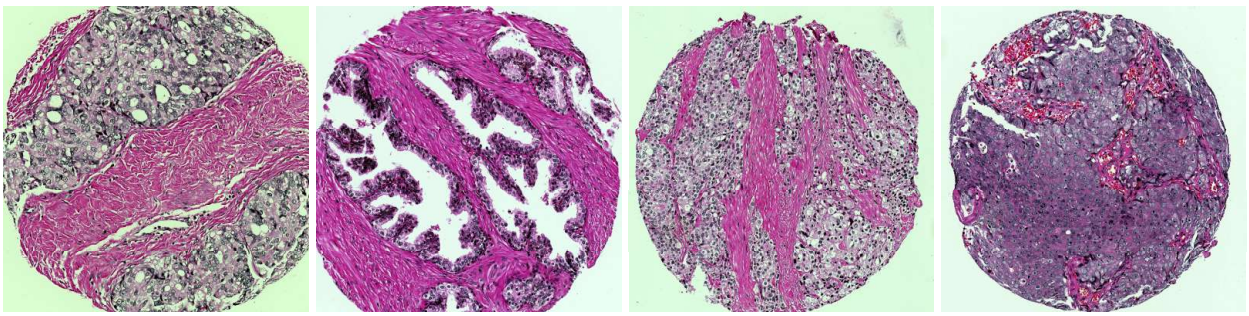
labels were created with the help of pathologists. In Figure **3-4** , we can see an example of our data set with their respective labels.

3.3.2 Preprocessing

Before entering the training and validation stage of the neural network, we submitted the database to two pre-processes. The first pre-processes consist of the white balance technique [47] that allows us to improve the image's contrast (Figure **3-5**b). After the technique is applied, we scale the image in the range of [0, 1]. The goal of scaling the image is to change the pixel values to use a common scale without altering the differences between value ranges or losing information. Also, this step is important for neural network models to improve the training process.



(a) Original Images



(b) White Balance Process

Figure **3-5**: Processing of the database using the white balance technique.

3.3.3 Network Architecture

The MobileNet [48] architecture was used with a perceptual field of 124×124 pixels. The use of predefined networks saves us from going through the iterative process involved in designing a neural network. In addition, this network has the fewest number of parameters (compared to the other predefined networks) and has been used to address image classification tasks [49]. In this work, we adapt the network to classify whether or not the central pixel of any given patch of 124×124 pixels in size corresponds to the stroma. However, there is no central pixel in a 124×124 pixel window, but rather a 2×2 square as shown in Figure **3-6**. We choose the pixel located in the lower right corner of the 2×2 windows as the central pixel of the perceptual field.

When we talk about a central pixel, we refer to a value of the matrix. This value is usually a half pixel. However, this is true when the dimension of the patch is an odd number. In cases where they are an even number, the central pixel is not extracted so directly. The neural network creates a small 2×2 square where you must select the central pixel using one of its sides, as explained above. The selection of the pixel is given by analyzing with a specific example, where our neural network selects the lower right pixel.

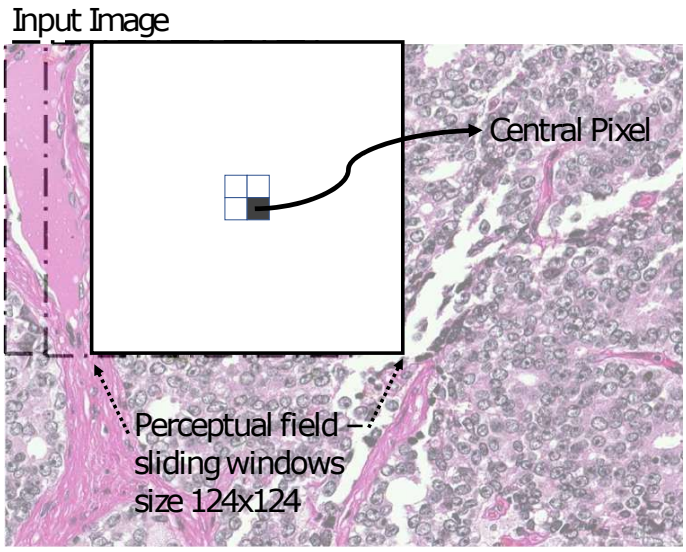


Figure **3-6**: CNN Perceptual Field. The dotted boxes indicate that the window slides across the image, overlapping each window and getting a label at the end.

3.3.4 CNN Training and Validation

For network training, we use the Transfer Learning technique. This technique allows us to use the characteristics (weights) of a network already trained with a very large data set to take advantage of them in our classification task. The trained weights correspond to the ImageNet database.

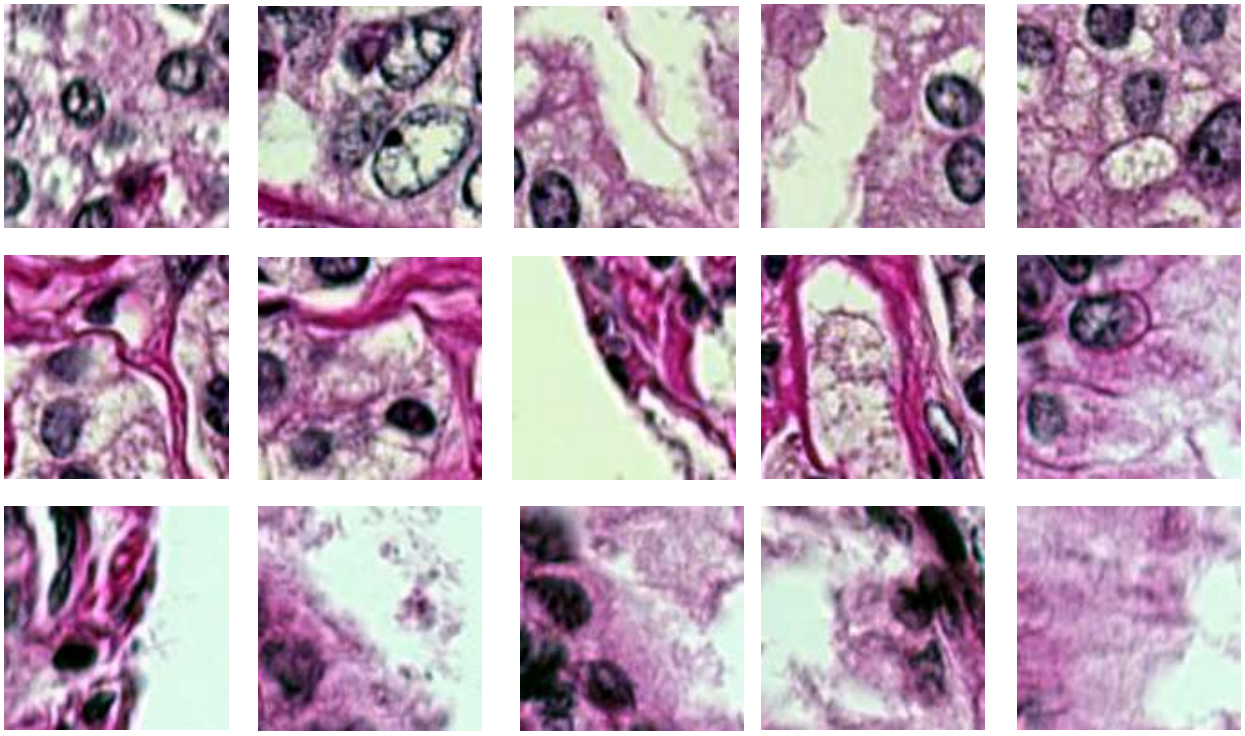
For the training set, we used 3,200 patches (1,600 background class patches and 1,600 Stroma class patches), and for the test set, we used 800 patches (400 background class patches and 400 Stroma class patches).In Figure **3-7**, we can see some examples of the patches used for each class of the training and test sets.

The training parameters and the Transfer Learning application were taken according to the Keras documentation[1].

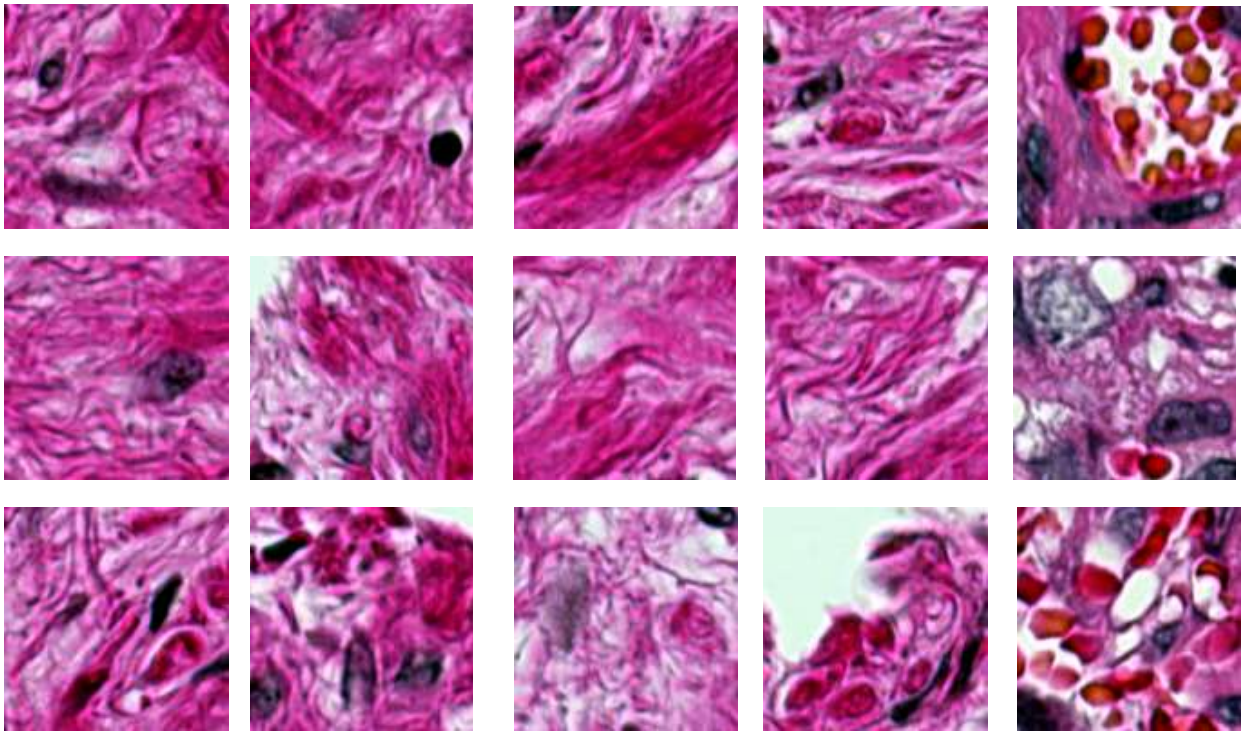
To improve the training performance of our neural network, we employ fine-tuning, which consists of retraining the CNN with a lower learning rate with a smaller number of epochs. In Table **3-1** we find the training parameters of each part (CNN Training and CNN Fine Tuning)

During training, our model achieved an accuracy of 0.9399 and a loss function of 0.2 after five epochs. The training was carried out using the Google Collaboratory platform in a virtual

[1] $https://keras.io/guides/transfer_learning/$



(a) Example of Background patches



(b) Example of Stroma patches

Figure **3-7**: Example of the patches of each class (background and stroma). Each patch is 124x124 pixels in size.

CNN Training Parameters		
Parameters	**CNN Training Part 1**	**CNN Fine-tuning Part 2**
Loss Function	*Binary Cross Entropy*	*Binary Cross Entropy*
Optimizer	*Adam*	*Adam*
Learning rate	*-2*	*-5*
Epoch	*20*	*5*

Table **3-1**: Parameter table for both trainings of the convolutional neural network

machine with 12Gb of RAM, 64Gb of hard disk, and a Tesla K80 GPU. The algorithms were coded in Python using the Keras, Tensorflow, Opencv, and Numpy libraries.

3.3.5 Inference - Proposed Algorithm

Once the network has been trained, the next step is to segment an image and compare the result with the label image. To carry out the segmentation, we have to use a window the size of the patches and slide it over the entire image. Each patch will be superposed, generating a label in each patch evaluation by the model.

However, as discussed in the previous section, this process is time-consuming since each image of size N×M pixels requires N*M evaluations of the network to be segmented. For this, we proposed to make a change in the segmentation methodology. This consists of an incremental solution providing a subset of randomly chosen sliding windows. The segmented image is formed by evaluating said subset with the CNN and assuming that the labels obtained from said evaluation affect the label of the central pixel and an area (or neighborhood) that includes and surrounds the pixel.

For the implementation of our method, we divide the algorithm into 2 parts:

Creation of the point cloud

As mentioned above, we randomly create sliding windows to test each window with the CNN. To do this, we transform the RGB image to L*a*b to use luminance as the base image. To this component, we apply the scale-space algorithm. Since the images in our database have structures at different scales, it is necessary to use this algorithm to make the image's characteristics invariant to scale. The next step is randomly selecting a percentage of pixels following a uniform distribution. Each pixel corresponds to the central pixel of each

sliding window. We call this selection of points the sampling rate. In Figure **3-8**, we can see how point clouds are generated with different sampling rates.

Point Cloud Evaluation

We use the generated point cloud to locate the central pixels of each window in the image. We then evaluate those windows and accumulate the result into a vector. Finally, this vector is transformed into a matrix corresponding to the segmented image.

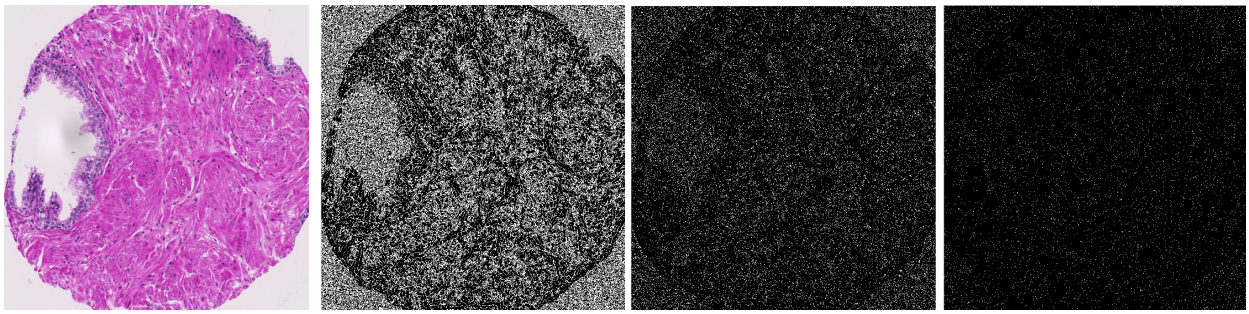
3.4 Results

To analyze if our method allowed us to segment the image faster and with a minimum of information loss, we used a series of analyzes and metrics to evaluate it. First, we take as the main basis a segmentation with a sampling rate of 100%; that is, all the image pixels were used. We then use sampling rates of 50%, 25%, 12.5%, 6.25%, 3.125%, and 1.625%. We compared each sampling rate to 100% to see how similar these resulting images were. In table **3-2**, we can see the confusion matrices for each sampling rate. Note that the percentages on the main diagonal correspond to how much the sampling rate resembled the main sampling rate. The inverted diagonal corresponds to the error percentages compared with the main sampling rate.

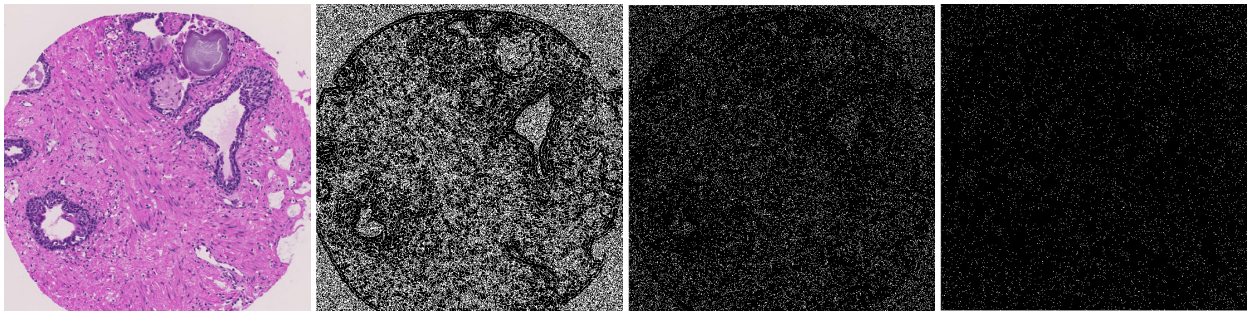
As we increase the sampling rate, the True Positive and True Negative values (The main diagonal of the confusion matrix) become more like the original. The range of these values is between 98% (the maximum) and 94.5% (the minimum). If we use only 1.625% of the image pixels, we can obtain 94.5% True Positives and 96.5% True Negatives from the segmentation performed with the 100% sampling rate.

To reinforce our results, we present three segmentation metrics widely used in the literature: precision, sensitivity, and the Dice index. The Dice index is a metric that measures, on a scale from 0 to 1, the level of similarity between a binary image A and another B (a value of 0 indicates that they are different, and a value of 1 indicates that they are the same). The precision quantifies the number of positive class predictions that belong to the positive class. Recall quantifies the number of positive class predictions made from all positive examples in the data set. Table **3-3** shows the results. These percentages correspond to how close each metric was to the metric generated by the main sampling rate.

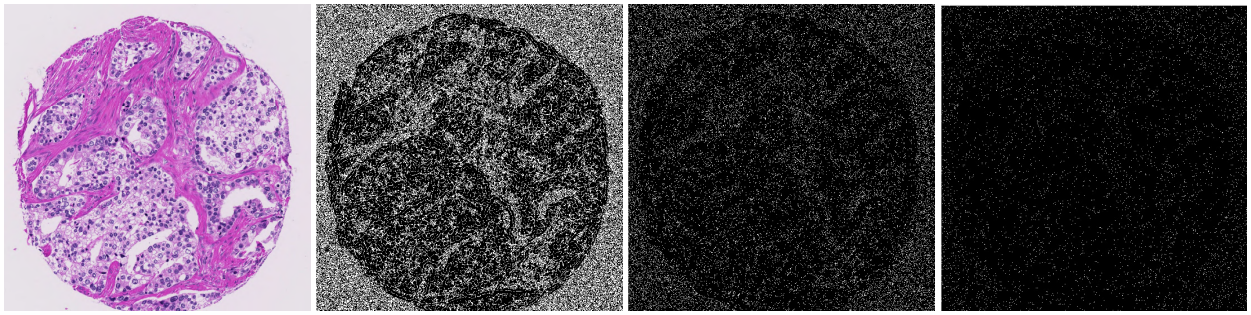
Using the definition of true positive (TP), false positive (FP) and false negative (FN), the



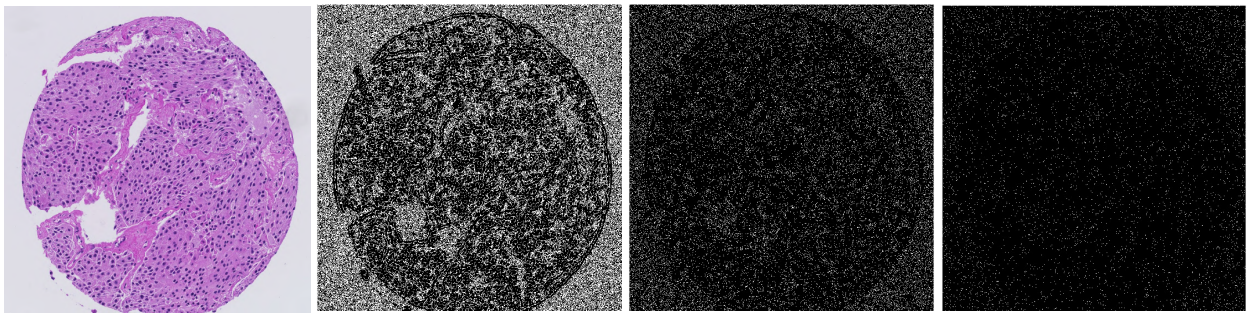
(a) Point cloud for Benign tissue



(b) Point cloud for Gleason 3 tissue



(c) Point cloud for Gleason 4 tissue



(d) Point cloud for Gleason 5 tissue

Figure **3-8**: Example of point clouds generated by different sampling rates. The first column corresponds to the original image. The second column corresponds to the point cloud with a sampling rate of 50%. The third column corresponds to the point cloud with a sampling rate of 12.5% and the last column corresponds to the point cloud with a sampling rate of 3.125%.

		Predicted	
		0	**1**
Actual	**0**	97.2%	2.8%
	1	2.0%	98.0%

(a) Confusion Matrix Rate sampling 50%

		Predicted	
		0	**1**
Actual	**0**	97.2%	2.8%
	1	2.1%	97.9%

(b) Confusion Matrix Rate sampling 25%

		Predicted	
		0	**1**
Actual	**0**	96.8%	3.2%
	1	2.8%	97.2%

(c) Confusion Matrix Rate sampling 12.5%

		Predicted	
		0	**1**
Actual	**0**	96.7%	3.3%
	1	3.4%	96.6%

(d) Confusion Matrix Rate sampling 6.25%

		Predicted	
		0	**1**
Actual	**0**	96.4%	3.6%
	1	4.5%	95.5%

(e) Confusion Matrix Rate sampling 3.125%

		Predicted	
		0	**1**
Actual	**0**	96.3%	3.7%
	1	5.5%	94.5%

(f) Confusion Matrix Rate sampling 1.5625%

Table **3-2**: Confusion matrices for each sampling rate. Note that the difference between the sampling rate of 50% and that of 1.5625% have a difference of less than 4% in each cell.

equations of the metrics used can be written as follows:

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{3-2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3-3}$$

$$Recall = \frac{TP}{TP + FN} \tag{3-4}$$

	Dice Index	**Precision**	**Recall**
50% sampling rate	97.59	97.20	97.98
25% sampling rate	97.54	97.20	97.88
12.5% sampling rate	96.99	96.80	97.88
6.25% sampling rate	96.65	96.70	96.60
3.125% sampling rate	95.96	96.40	95.54
1.5625% sampling rate	95.44	96.30	94.59

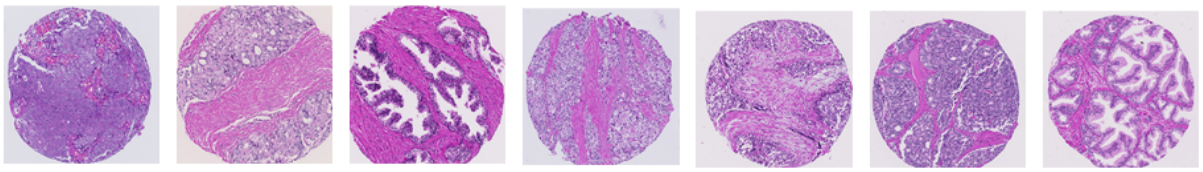
Table **3-3**: Values of the metrics used in each sampling rate. Note that there is less than 4% loss of information as we use a smaller value of the sampling rate.

As fewer pixels are used for segmentation, the time required for the proposed method to give us a labeled image decreases markedly. In Table **3-4**, we can see the average time required for each sampling rate.

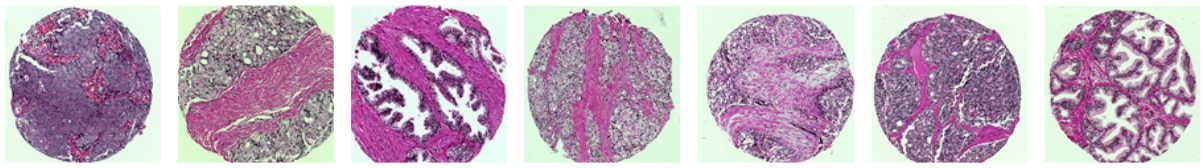
Sampling rate value (%)	**Time**
100	>24h 0' 0"
50	1h 52' 18"
25	0h 42' 36"
12.5	0h 23' 24"
6.25	0h 12' 36"
3.125	0h 6' 36"
1.5625	0h 4' 08"

Table **3-4**: Average time required to segment an image using different sampling rates.

Finally, in Figure **3-9** we can see some examples of segmentation with the different sampling rates.



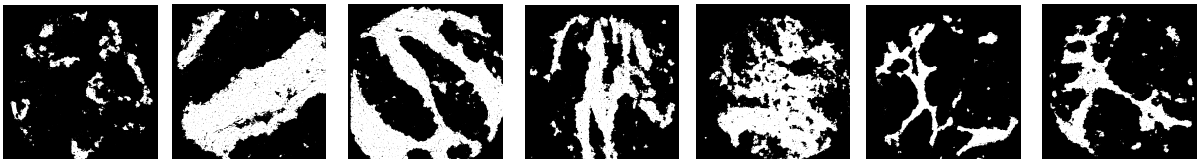
(a) Original Image.



(b) Pre-processed image.

(c) Ground truth.

(d) Stroma Segmentation using exhaustive segmentation (all the pixels in the image).



(e) Stroma Segmentation using only 25% of the pixels in the image).

(f) Stroma Segmentation using only 6.25% of the pixels in the image).

(g) Stroma Segmentation using only 1.5625% of the pixels in the image).

Figure **3-9**: Results of the stages of our proposed methodology

3.5 Conclusion

The designed method allowed us to reduce the time required by a convolutional neural network to segment an image. We use different sampling rates of 50%, 25%, 12.5%, 6.25%, 3.125%, and 1.5625% to analyze our algorithm's convergence through confusion matrices. Additionally, we use metrics such as the Dice index, recall, and precision to validate that the method converges experimentally. Using only 1,625% of the pixels of an image, we obtained 94% similarity concerning the 100% sampling rate (using all the pixels) in an average time of 5 minutes; that is, we managed to reduce the time significantly considerable without major loss of information.

4 Location of suspicious areas using CNNs

In the previous section, we designed a methodology (Transfer Learning with the MobileNet network) that allowed us to segment muscle tissue in images of prostate tissue. We were able to optimize the segmentation (inference) time of our convolutional neural network only using a percentage of image pixels, where we managed to reduce the time from 24 hours to just 5 minutes with a very low loss of information (around 4%). However, in the state of the art, a very important task carried out with neural networks in medical imaging is to locate tissue abnormalities. Using the results of chapter 2, in this chapter, we will implement a methodology that allows us to locate suspicious areas in images of prostate tissue. For this, we will design our own neural network and train it from scratch. This will reduce the inference complexity and also the time needed to segment the image. This chapter is organized as follows. In section 1 we will make an introduction to the different existing neural architectures in the literature that are used in histopathological images and the problem that can be extrapolated from these deep networks. In section 2 we will show our proposal convolutional neural network to locate suspicious tissue areas. In section 3 we will present the results obtained and finally, in section 4 we will give the conclusion of the chapter.

4.1 Introduction

It is common to see authors employ predefined neural architectures to solve computer vision problems in the state-of-the-art [50]. By using predefined models that have been trained on a large data set, the training process becomes much more manageable. Among the benefits that exist when using this methodology, we can mention the most significant:

- **Few training data:** As previously trained weights are used, few images can be used to adapt the convolutional model to the new data set. This process is known as Transfer Learning.
- **High Network Performance:** When the neural network is retrained with the new

data, the model starts with good performance (good accuracy, little overfitting, few parameters to tune, etc.). For this reason, few periods and small learning rates are used.

In the following articles, we can see how this methodology generates good results using histopathological images. Hameed et al. [51] present a deep learning approach for breast tissue classification on histopathological images. The database used consisted of 845 images of size 1278×760 pixels extracted from 544 Whole-Slide Images. The classes used were images with carcinoma (437) and images without carcinoma (408). They used the Transfer Learning technique with the VGG19 and VGG16 [52] networks for training. The perceptual field of the network was $224 \times 224 \times 3$ pixels (the image was scaled to 1278×760 pixels). The procedure used these patches to classify them into one of the two classes described above. The results obtained by the authors were 95.29% in the F1 score, 97.73% in the recall, and 95.29% in the precision.

Sitaula et al. [53] also used the Transfer Learning technique with the VGG16 pre-trained neural network to classify histopathological images of breast tissue. The authors used the BACH database of 400 images of 2048×1536 pixels. The perceptual field of CNN is $224 \times 224 \times 3$ pixels. However, the authors did not scale the image but instead used extracted patches of 224×224 size as input to the neural network. It is required to classify the images into 4 different classes: normal, benign, in situ, invasive. The inference process consisted of using the patches as features of an image and using a support vector engine to classify these features into one of 4 classes. In the normal class, the results were 90% in the precision, recall, and F1-score. The results were 91%, 89%, and 90% for the benign class. The results were 94%, 94%, and 94% for the in situ class. Finally, the results were 93%, 96%, and 94% in the invasive class.

In these examples, we saw how they used the VGG16 and VGG19 neural architecture (Figure **4-1** [1]) to classify images of breast tissue. However, this pre-trained network contains around 138 million parameters. Therefore, the number of parameters affects the training and inference time, and it is required to use hardware with greater capacity to handle this type of network.

This is not the only architecture used in histopathological image classification. For example, Han et al. [54] used the AlexNet neural network (Figure **4-2**[2]) to localize areas with a high degree of cancer in images of prostate tissue. They used 268 WSI images from which 960×960 pixel images were extracted. Each portion of the image corresponds to the regions

[1] $https://www.researchgate.net/figure/Fig-A1-The-standard-VGG-16-network-architecture-as-proposed-in-32-Note-that-only_fig3_322512435$

[2] $https://neurohive.io/en/popular-networks/alexnet-imagenet-classification-with-deep-convolutional-neural-networks/$

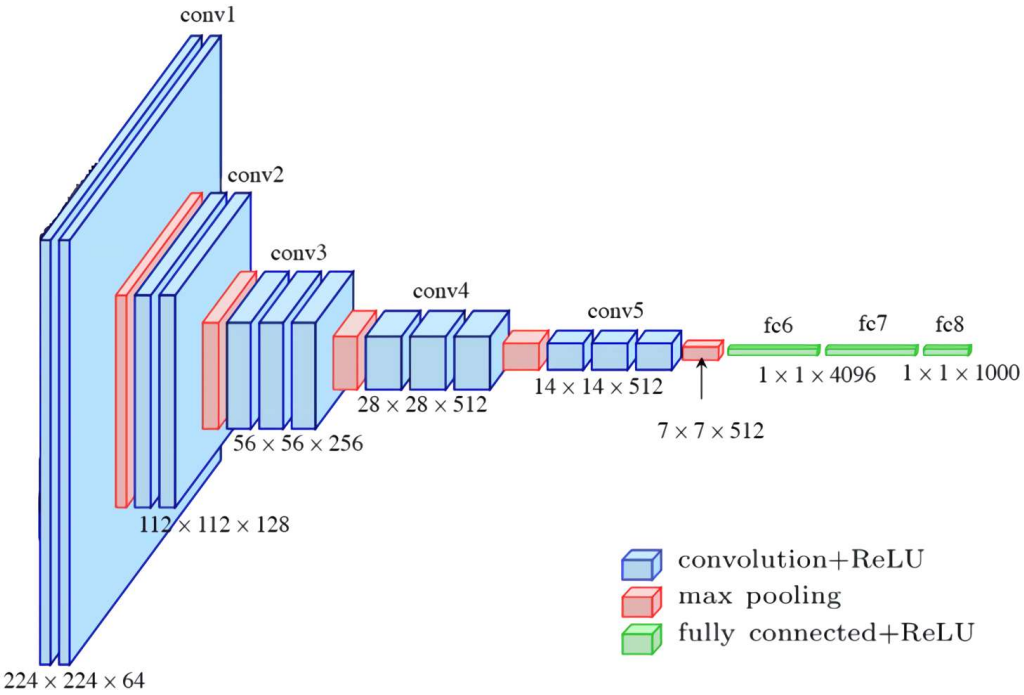


Figure **4-1**: Basic architecture of the VGG16 neural network. This network contains about 138 million parameters. VGG19 architecture is similar to VGG16, the difference is that VGG19 contains three additional convolutional layers.

of interest of the investigation. Two labels were used: cancer class and benign class. The training process consisted of extracting nuclei and Lumen features from the images. They then used the pre-trained model to detect the zones with the highest nuclei concentration and classify this characteristic in the cancer class and the rest of the zones (highest concentration of Lumen) in the benign class.

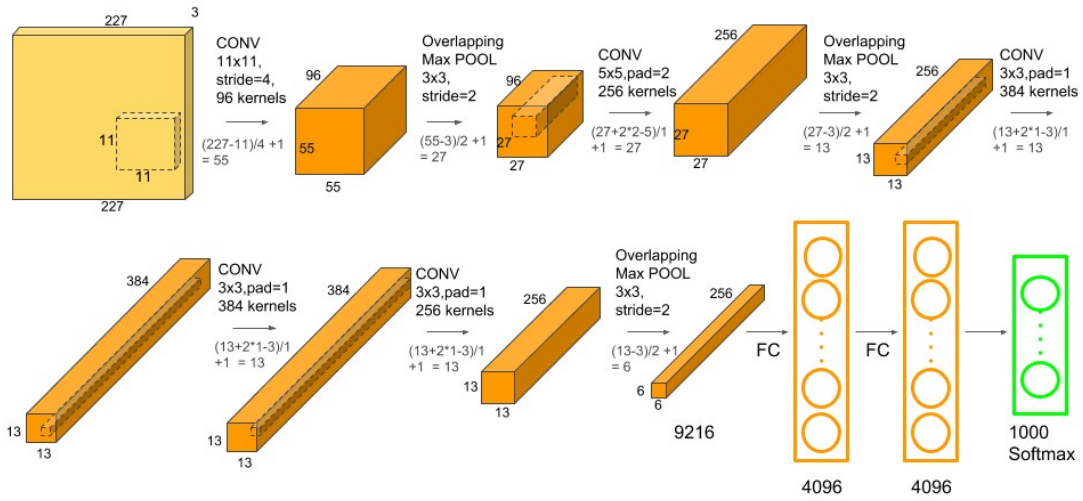


Figure **4-2**: Basic architecture of the AlexNet neural network. The neural network contains 61 million parameters.

As a last example, Wang et al [55] develop an automated system that allows the recognition of tumor areas for pathology images of lung cancer. They used 27 WSI images from which patches of 300x300 pixels were extracted. In total, 5,344 patches were extracted, of which 2,139 corresponded to the benign class, 2,475 to the tumor class, and 750 to the white background class. The Inception V3 architecture (Figure **4-3**[3]) with a 300×300 perceptual field and pre-trained weights in ImageNet was used to train the CNN model. The overall prediction accuracy of the CNN model on the test set was 89.8%; accuracy was 88.1% for tumor patches and 93.5% for benign patches.

As we can see in the previous examples, pre-trained neural architectures to solve histopathological image classification and segmentation problems have a good performance index. However, this methodology offers good performance at a high price: the weight of the network and the high number of parameters. Thanks to the designed method in chapter 2, we can create more data by dividing the image into small patches. This implies that it is not necessary

[3] $https://www.researchgate.net/figure/Convolutional-neural-network-architecture-Inception-v3-used-in-this-study-Inception-v3_fig3_328775405$

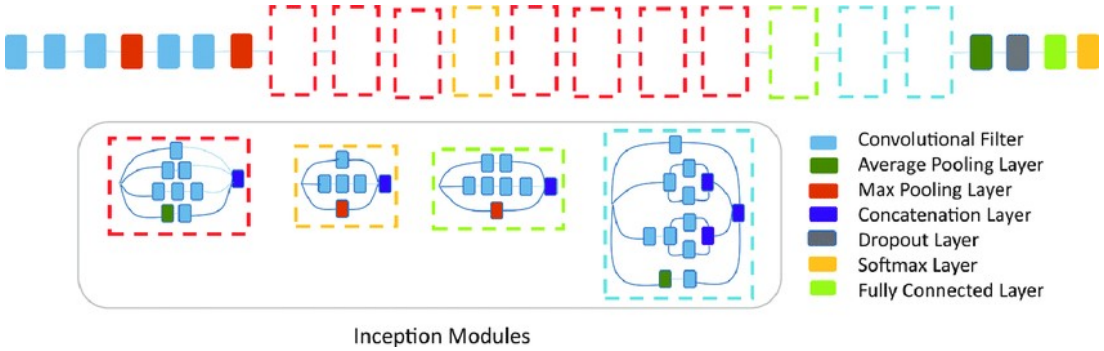


Figure **4-3**: Basic architecture of the Inception V3 neural network. The neural network contains 24 million parameters.

to use the Transfer Learning technique. Therefore, we can simplify the problem by creating a convolutional neural network with much fewer parameters and good performance, as in the examples cited. In section 2, we will present our new methodology to locate suspicious areas in images of prostate cancer.

4.2 Proposed methodology

In state-of-the-art, they expose different ways to optimize a neural network, from better training forms [56, 57] to optimizing the inference time [58, 59]. In this document, we select one of the elementary techniques that optimize the processing time: reduce the size of the image. To use this technique, we make an assumption: that the characteristics of the image are maintained, that is, that the loss of information, when going from the original scale to a smaller one, is low. Therefore, the perceptual field is also reduced since we are not using the original image (compared to how it was done in chapter 2). Due to the simplifications that were made, it was proposed to divide the methodology into two separate processes that are executed sequentially:

- **Stroma segmentation.**
- **Location of abnormal areas.**

4.2.1 DataBase

The database corresponds to the same database that was used in chapter 2 of this document (See Database section of Chapter 2)

4.2.2 Preprocessing

Three processing techniques were used to have a data set as uniform as possible. We first applied the staining normalization technique [60]. This allows us to reduce the variance of the staining that exists between prostate images. Second, we use the white balance technique to increase the contrast of the image structures: stroma, nuclei, glands, and lumen (it is the same technique used in the processing stage of chapter 2). Finally, we scaled the images from 3100x3100 pixels to 752×752 pixels. In Figure **4-4**, we observe these techniques.

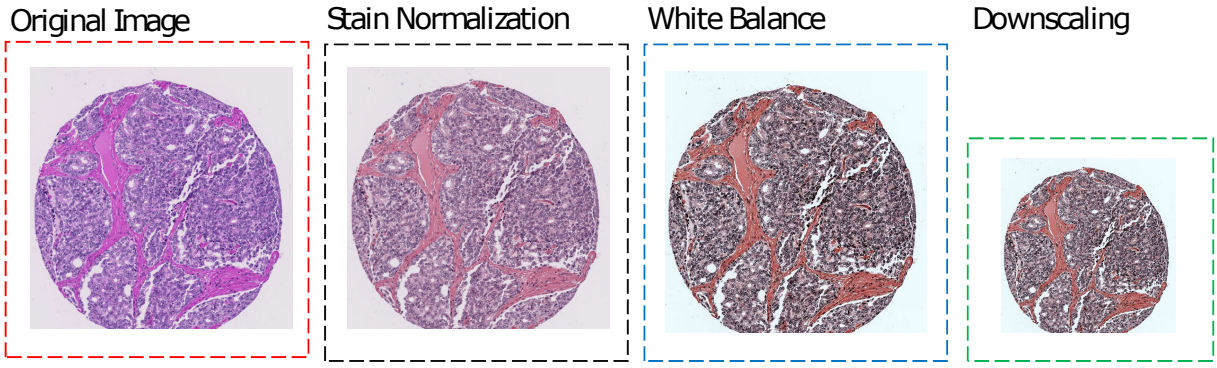


Figure **4-4**: Database preprocessing. The image corresponds to a Gleason 4 tissue.

4.2.3 Architecture Network

The architecture of the convolutional neural network can be seen in Figure **4-5**. The feature extraction process consists of three convolutional layers and two maximal pooling layers. The classification process consists of 32 neurons with a dropout of 0.5, followed by the network's output containing two neurons. The perceptual field of the network, that is, the size of the input to the network, is $32 \times 32 \times 3$. The size was selected by visual inspection, where the stroma and epithelial tissues were analyzed for the smallest dimension. It may contain enough information for the neural network to work well. Selection criteria were based on the amount of muscle tissue and nuclei that could cover the perceptual field of the network and the number of cells present in the cytoplasm that could cover the same perceptual field of the network.

4.2.4 Stroma segmentation

Using the network in Figure **4-5**, we proposed segmenting the stroma of the images as done in chapter 2. The objective is to eliminate pixels that are not required for the location of

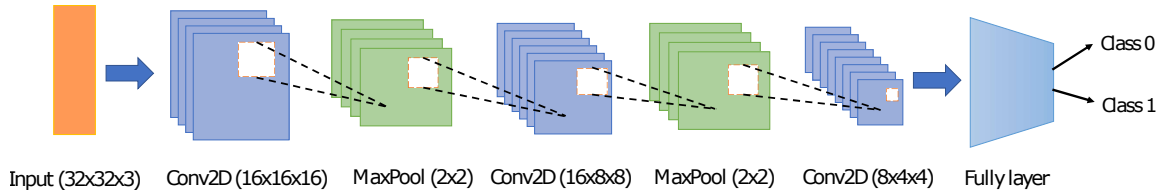


Figure **4-5**: Convolutional neural network proposal. Our CNN contains only 47242 trainable parameters.

anomalies (second stage). As the convolutional neural network and the perceptual field is small, the required inference time is low.

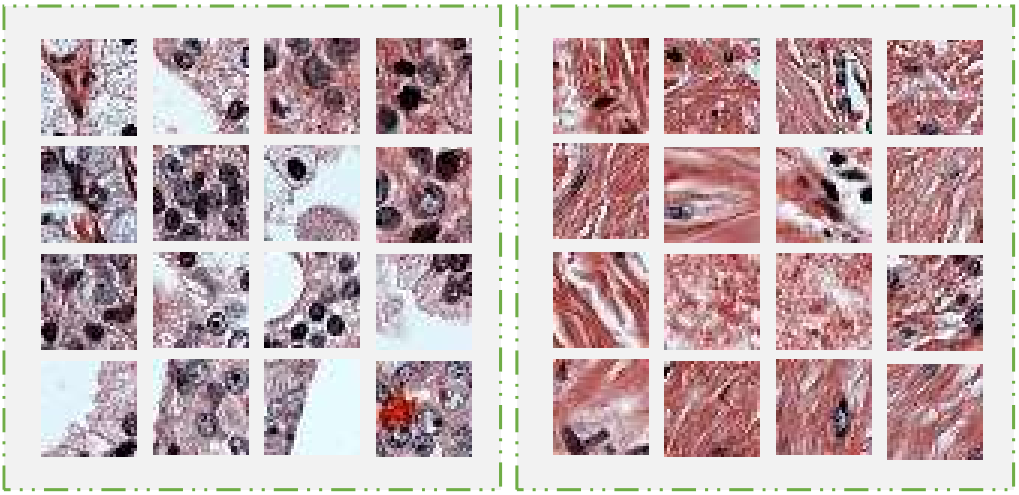
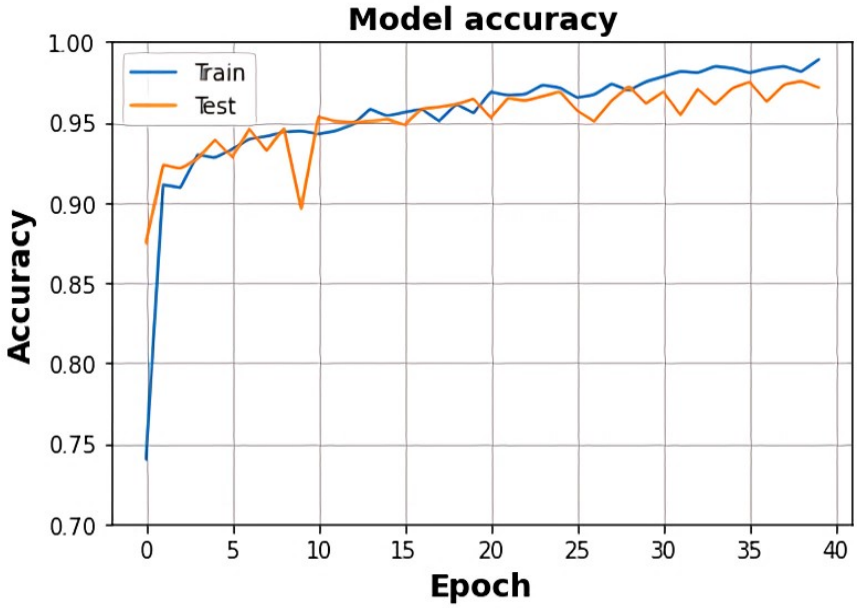


Figure **4-6**: Example of the training and test images. The box on the left is the patches of the background class and the box on the right is the patches of the stroma class.

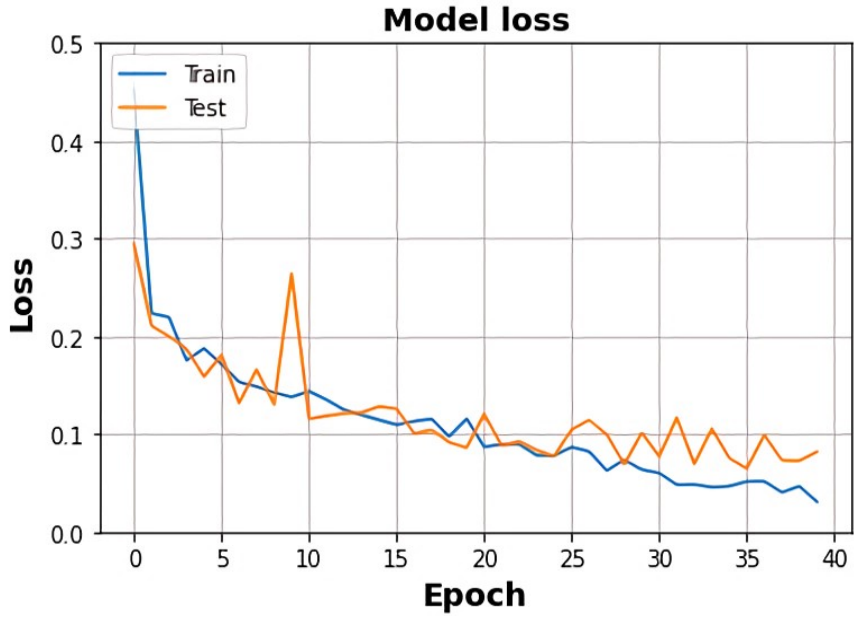
To train the network, 1500 patches of $32 \times 32 \times 3$ in two different classes were used: Background Class and Stroma Class, and for model validation, 500 patches were used for each class (See example Figure **4-6**). To train our network, we use 40 epochs with a batch size of 8. We use the Adam optimizer with a training rate of 0.0001. In Figure **4-7**, we can see the result of the network training

4.2.5 Location of abnormal areas

After segmenting the stroma of the images, new patches are extracted to enter the same neuronal model whose function is to detect suspicious areas. The importance of the first segmentation is that it allows us to focus on the most relevant tissues. Removing a good percentage of the muscular tissue from the image will reduce the possible false positives in the inference since the network will not be confused with the stroma tissue. The classes that



(a) CNN Accuracy plot



(b) CNN Loss plot

Figure **4-7**: Plots of neural network training corresponding to Stroma segmentation

were used were: benign class and malignant class. In training, 4000 patches were used for each class and 900 patches for model validation (See example Figure **4-8**). The neural model has the same training parameters described in the first stage.In figure **4-9**, we observe the training graphs.

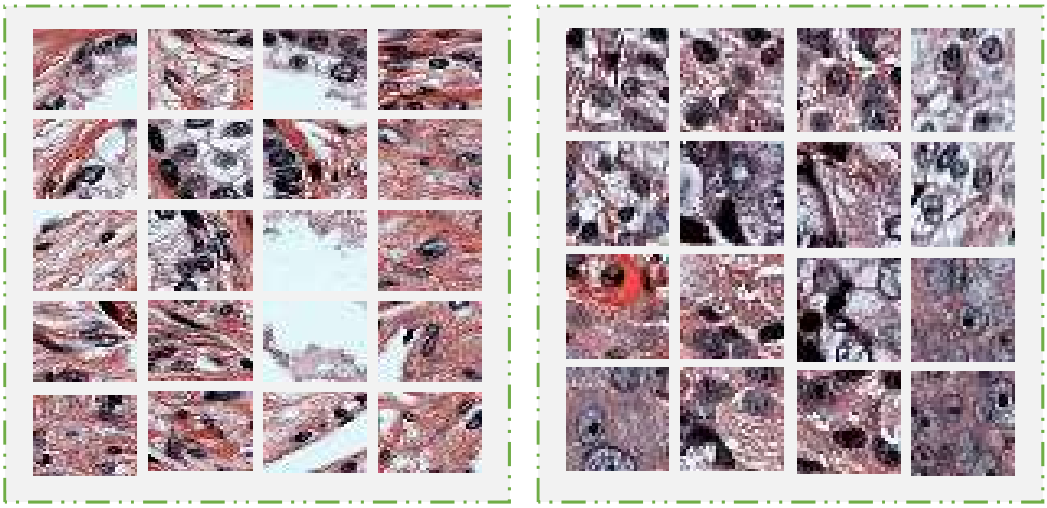


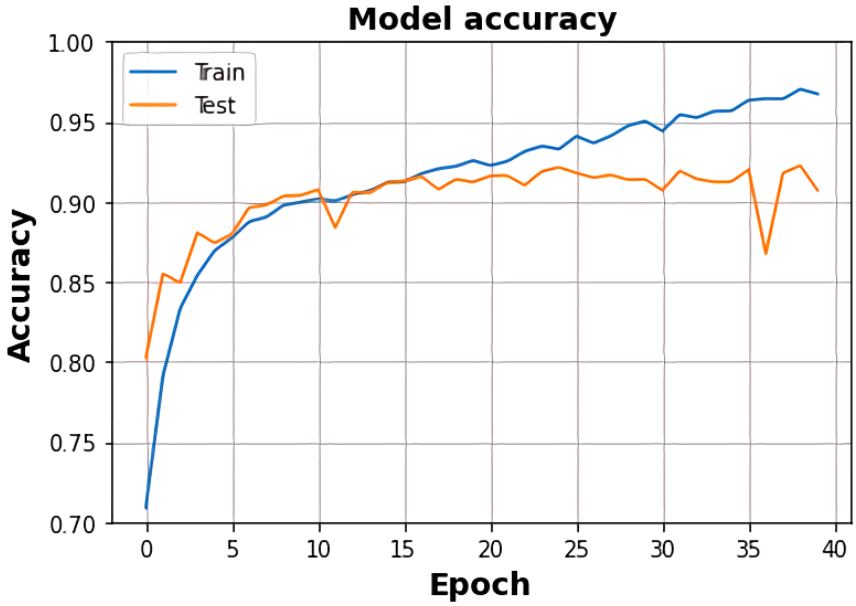
Figure **4-8**: Example of the training and test images. The box on the left is the patches of the benign class and the box on the right is the patches of the malignant class.

4.3 Results

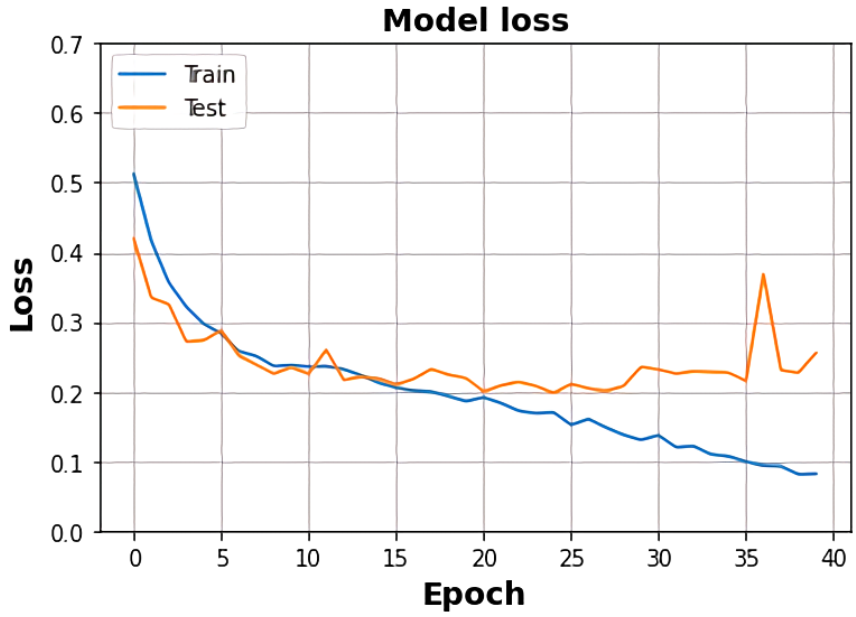
4.3.1 Stroma Segmentation

The neural network showed a performance of 96.7%; that is, of the 1000 images used for the test, 967 were classified correctly and 33 incorrectly. In Stroma segmentation, the neural network took an average of 30 seconds. In Figure **4-10**, we see the result of some classifications.

In some images, the presence of stroma is observed. This is because epithelial and stroma tissues have similar texture characteristics, causing false positives in segmentation. However, CNN manages to discriminate a good percentage of muscle tissue from the image's background. Using this result, we can use the second neural network to locate areas with a high degree of cancer.



(a) CNN Accuracy plot



(b) CNN Loss plot

Figure **4-9**: Plots of neural network training corresponding to location of abnormal areas

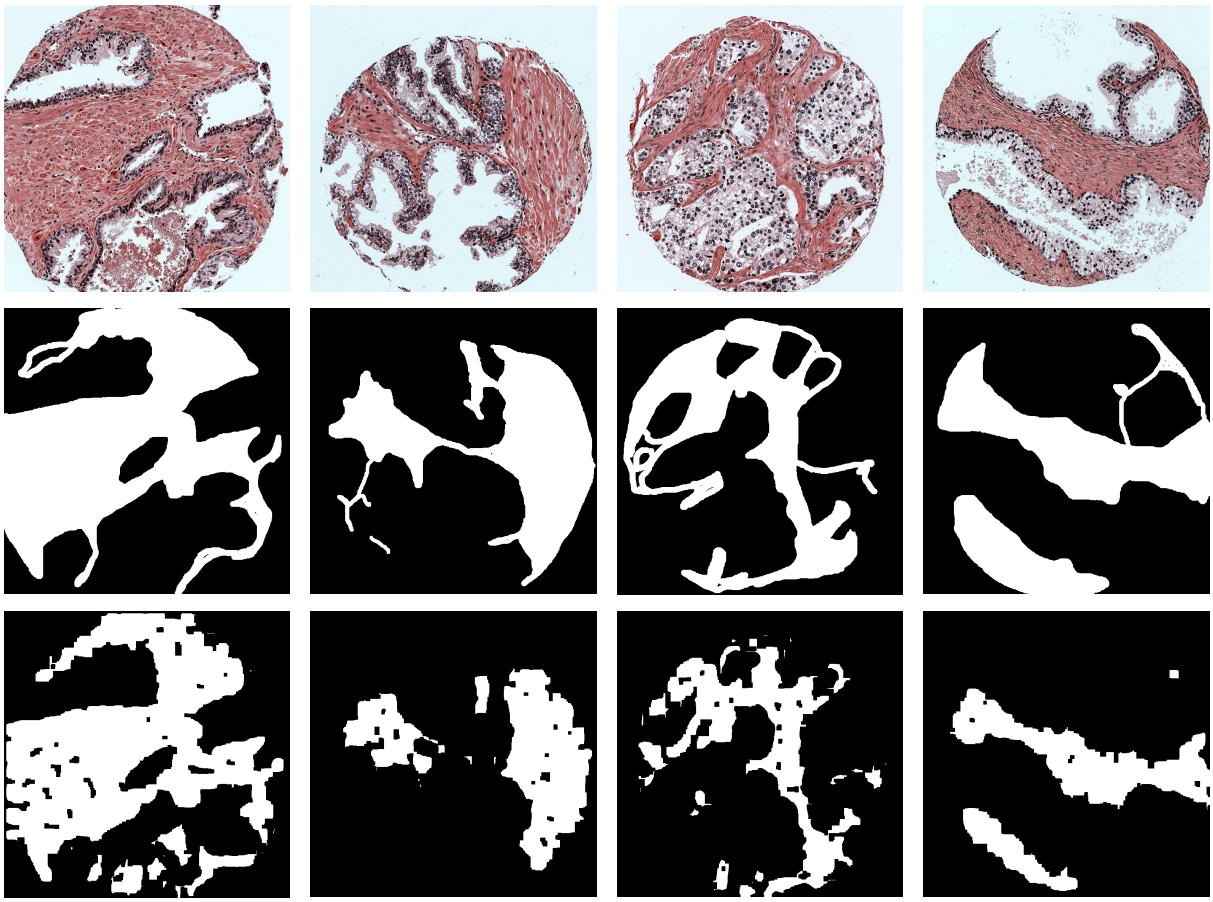


Figure **4-10**: Example of the result of the segmentation in images. The first row corresponds to the processed image, the second row to the label of the stroma and in the last row the image segmented by CNN.

4.3.2 Location of abnormal areas

The neural network showed a performance of 90.72%; that is, of the 1,800 images used for the test, 1,633 were correctly classified and 167 incorrectly. The average segmentation time was 45 seconds. The segmentation process at this stage consisted of training the same CNN but with the new labels. The result is a black and white image with the areas that the network considers suspicious. However, as mentioned above, stroma tissues and glands have similar characteristics, causing segmentation to have poor performance. For this reason, the same CNN neural network was trained for two different tasks. When we combine the information from the first neural network (stroma segmentation) with the information from the second neural network (location of suspicious areas), the result is a binary image where the maximum value corresponds to cancer and the minimum value to benign tissue. This allows mitigating the false positives generated in the segmentation. Table **4-1** shows the operation performed with the two segmentation. Additionally, we refine the resulting image with a morphological operation to eliminate those small areas that do not represent a suspicious zone. In Figure **4-11**, we show the final results of the location of the malignant tumor areas.

Location of abnormal area	**Stroma Segmentation**	**Result**
1	1	0
1	0	1
0	1	0
0	0	0

Table **4-1**: True table. The values in the table correspond to the pixel values, where 1 equals the white pixel and 0 the black pixel.

4.4 Conclusion

The proposed design to locate suspicious areas has many advantages over the examples postulated at the beginning of the chapter. The neural network has a reduced number of parameters and has performance comparable to that of predefined networks. The strategy of using two CNNs to perform a task allows us to control the behavior of the neural network since, being such a small network, we can know what characteristics it is using to segment. Additionally, the time required for training and inference is reduced (less than one minute), allowing different iterative tests to be carried out to find a model with better performance. We show experimentally that it is not necessary to use such deep networks (over-sizing of the problem); it is enough to make some assumptions at the processing level to design much smaller models with the same level of precision.

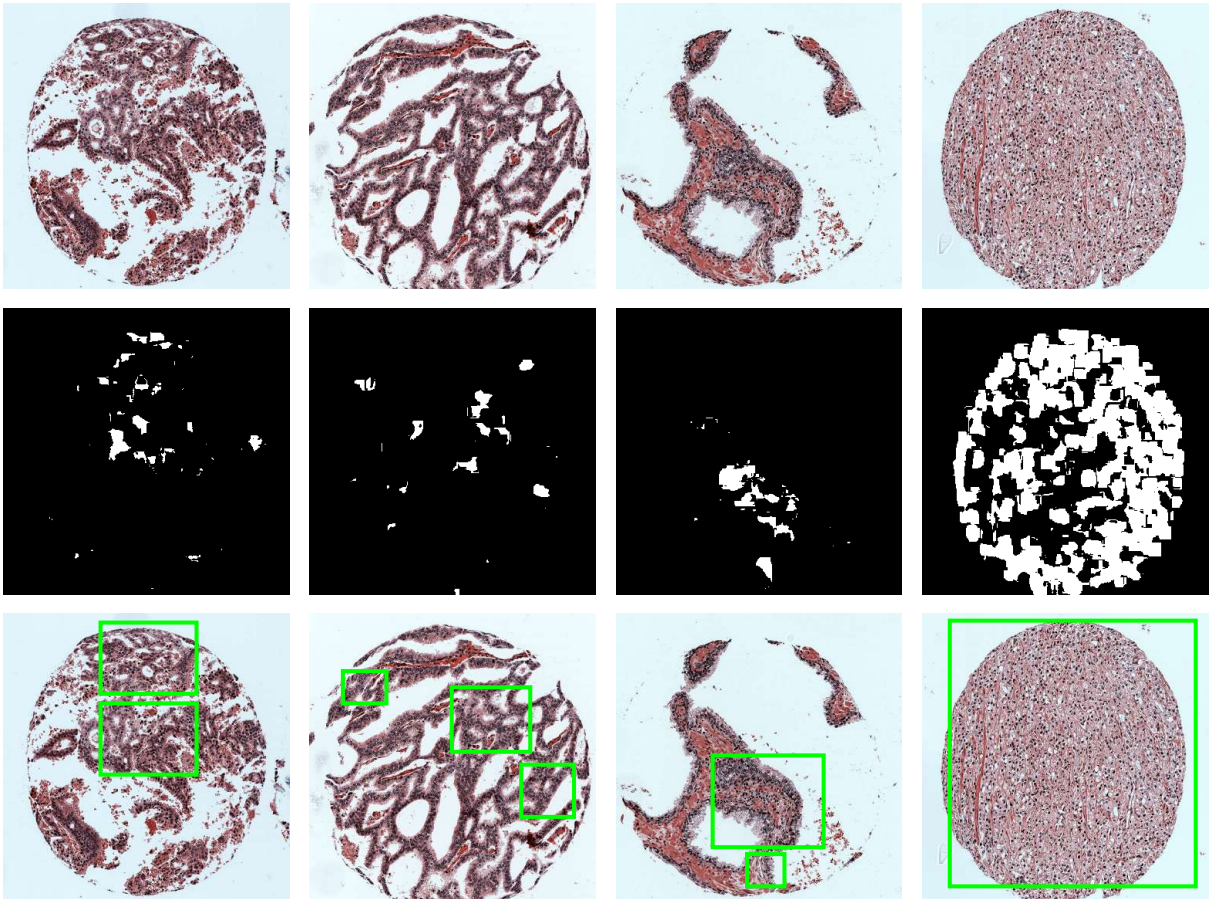


Figure **4-11**: Example of the location of areas with a high incidence of cancer. The first row corresponds to the original image; the second row refers to the result of applying Table **4-1** with the results of the two CNNs, and the last row corresponds to the location of the tissue.

5 Conclusions

In this document, we saw three types of tools that facilitate some of the tasks that a pathologist carries out to diagnose a patient. The first tool is a series of color transformations that improve the visualization of the structures present in the images of prostate tissue. We used 44 color transformations, of which 52.27% (23 transformations) were qualified as applicable by specialists. For instance, the white balance technique, a technique that had the highest rating by pathologists, allowed to increase the image's contrast, making the edges of the nuclei better defined. In addition, both the stroma and the epithelial tissue differences were achieved much better than using the original image.

The second tool is an incremental algorithm that allows segmenting images in a pixel-based manner based on CNNs faster, without considerable loss of information. The task consisted of segmenting the stroma of prostate tissue, using a database of 100 images in multi-array format. In addition, the selected neural model was a pre-trained network called MobileNet. The designed algorithm significantly reduced the computational cost by using a randomly sampled subset of patches from the image, providing us with a way of reducing the time spent by computations from 24 hours of processing to just 5 minutes, with a minimal loss of information (less than 5%). In addition, this new methodology allows us to analyze many more images in a time lapse, facilitating the iterative process of validation and evaluation of future neural models.

Finally, the last tool uses the previous stage's results to locate suspicious areas in prostate cancer images.We designed a convolutional neural network that had a small number of parameters to reduce the computational cost of prediction. We hypothesize that using a pre-trained neural network such as AlexNet or VGG16 produces an oversized model for the task to be solved. Our network performs two tasks independently: segmenting the stroma and locating suspicious areas. Because finding suspicious areas generated a high number of false positives, we used the information from the stroma segmentation to discard those misclassified pixels. In the end, we obtained a tool capable of identifying suspicious areas with a reduced number of false positives.

6 Future work

In the first instance, we propose to analyze the assumptions made in chapter 3, where we affirm that, when scaling the original image, the present characteristics would not be modified, and there would be no loss of information. We also define a small perceptual field selected by simple visual inspection. Finally, it is proposed to carry out an experiment where the size of the network input is modified to see if the performance of the CNN varies.

During the development of the investigation, we observed that several histopathological tissues present a similar characteristic to prostate tissue. Therefore, we proposed to apply the transfer learning technique with our learned weights to see if our model adapts to new tissues.

Convolutional neural networks are not the only existing methodologies in deep learning. For example, there are neural models called fully convolutional networks like the U-Net specifically designed to perform segmentation. It is proposed to implement this type of neural network with our images and verify if there are advantages or if, on the contrary, our methodology continues to have excellent results.

Bibliography

[1] Justin Streicher, Brian Lee Meyerson, Vidhya Karivedu, and Abhinav Sidana. A review of optimal prostate biopsy: indications and techniques. *Therapeutic Advances in Urology*, 11:175628721987007, 2019.

[2] Eric Swanson and W. Dean Wallace. Histopathology Methods and Protocols. *Methods in Molecular Biology*, 1180:283–291, 2014.

[3] Ni Chen and Qiao Zhou. The evolving gleason grading system. *Chinese Journal of Cancer Research*, 28(1):58–64, 2016.

[4] John Murtagh. Gleason score.

[5] James A. Diao, Richard J. Chen, and Joseph C. Kvedar. Efficient cellular annotation of histopathology slides with real-time AI augmentation. *npj Digital Medicine*, 4(1):1–2, 2021.

[6] Shivang Naik, Anant Madabhushi, John Tomaszeweski, and Michael D. Feldman. A quantitative exploration of efficacy of gland morphology in prostate cancer grading. *Proceedings of the IEEE Annual Northeast Bioengineering Conference, NEBEC*, 08854:58–59, 2007.

[7] Scott Doyle, Mark Hwang, Kinsuk Shah, Anant Madabhushi, Michael Feldman, and John Tomaszeweski. Automated grading of prostate cancer using architectural and textural image features. *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*, pages 1284–1287, 2007.

[8] Feldman Naik, S., Doyle, S., Madabhushi, A., Tomaszewski, J. Automated nuclear and gland segmentation and gleason grading of prostate histology by integrating low-, high-level and domain specific information. *In: Special Workshop on Computational Histopathology (CHIP), in conjunction with 5th IEEE Internat. Symp. on Biomedical Imaging (*, pages 1–8, 2008.

[9] Kien Nguyen, Anil K. Jain, and Ronald L. Allen. Automated gland segmentation and classification for gleason grading of prostate tissue images. In *2010 20th International*

Conference on Pattern Recognition, pages 1497–1500, 2010.

[10] Kien Nguyen, Anindya Sarkar, and Anil K. Jain. Prostate cancer grading: Use of graph cut and spatial arrangement of nuclei. *IEEE Transactions on Medical Imaging*, 33(12):2254–2270, 2014.

[11] Kien Nguyen, Anil K. Jain, and Ronald L. Allen. Automated gland segmentation and classification for gleason grading of prostate tissue images. *Proceedings - International Conference on Pattern Recognition*, pages 1497–1500, 2010.

[12] Hong Kong. Proceedings of 2010 IEEE 17th International Conference on Image Processing CROSS-SECTIONAL MICROSCOPIC IMAGES Tian Xia Yizhou Yu Jing Hua Wayne State University University of Illinois at Urbana-Champaign. *Training*, pages 1057–1060, 2010.

[13] Hadi Rezaeilouyeh, Mohammad H. Mahoor, Francisco G. La Rosa, and Jun Jason Zhang. Prostate cancer detection and gleason grading of histological images using shearlet transform. *Conference Record - Asilomar Conference on Signals, Systems and Computers*, pages 268–272, 2013.

[14] Ali Tabesh, Mikhail Teverovskiy, Ho Yuen Pang, Vinay P. Kumar, David Verbel, Angeliki Kotsianti, and Olivier Saidi. Multifeature prostate cancer diagnosis and gleason grading of histological images. *IEEE Transactions on Medical Imaging*, 26(10):1366–1378, 2007.

[15] Ali Tabesh, Vinay P. Kumar, Ho-Yuen Pang, David Verbel, Angeliki Kotsianti, Mikhail Teverovskiy, and Olivier Saidi. Automated prostate cancer diagnosis and Gleason grading of tissue microarrays. *Medical Imaging 2005: Image Processing*, 5747:58, 2005.

[16] Siyamalan Manivannan, Wenqi Li, Jianguo Zhang, Emanuele Trucco, and Stephen J. McKenna. Structure Prediction for Gland Segmentation with Hand-Crafted and Deep Convolutional Features. *IEEE Transactions on Medical Imaging*, 37(1):210–221, 2018.

[17] Yan Xu, Yang Li, Yipei Wang, Mingyuan Liu, Yubo Fan, Maode Lai, and Eric I.Chao Chang. Gland Instance Segmentation Using Deep Multichannel Neural Networks. *IEEE Transactions on Biomedical Engineering*, 64(12):2901–2912, 2017.

[18] Philipp Kainz, Michael Pfeiffer, and Martin Urschler. Semantic Segmentation of Colon Glands with Deep Convolutional Neural Networks and Total Variation Segmentation. pages 1–15, 2015.

[19] Safiyeh Rezaei, Ali Emami, Nader Karimi, and Shadrokh Samavi. Gland Segmentation in Histopathological Images by Deep Neural Network. *2020 25th International Computer*

Conference, Computer Society of Iran, CSICC 2020, pages 1–5, 2020.

[20] Christophe Avenel, Anna Tolf, Anca Dragomir, and Ingrid B. Carlbom. Glandular Segmentation of Prostate Cancer: An Illustration of How the Choice of Histopathological Stain Is One Key to Success for Computational Pathology. *Frontiers in Bioengineering and Biotechnology*, 7(July):1–11, 2019.

[21] Hao Chen, Xiaojuan Qi, Lequan Yu, and Pheng Ann Heng. DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-December:2487–2496, 2016.

[22] Yuanping Zhou, Changqin Shi, Bingyan Lai, and Giorgos Jimenez. Contrast enhancement of medical images using a new version of the World Cup Optimization algorithm. *Quantitative Imaging in Medicine and Surgery*, 9(9):1528–1547, 2019.

[23] Yawu Li, Ning Li, Xiang Yu, Kai Huang, Ting Zheng, Xiaofeng Cheng, Shaoqun Zeng, and Xiuli Liu. Hematoxylin and eosin staining of intact tissues via delipidation and ultrasound. *Scientific Reports*, 8(1):1–8, 2018.

[24] Krishnan Nallaperumal, Muthukumar Subramanyam, Ravi Subban, Pasupathi Perumalsamy, Shashikala Durairaj, S. Gayathri Devi, and S. Selva Kumar. An analysis of suitable color space for visually plausible shadow-free scene reconstruction from single image. *2013 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2013*, (October 2014), 2013.

[25] Arash Abadpour. Color Image Processing Using Principal Component Analysis. (July):697, 2005.

[26] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Pearson Education, United States, 2008.

[27] Nicholas McCarthy, Padraig Cunningham, and Gillian Ohurley. The contribution of morphological features in the classification of prostate carcinoma in digital pathology images. *Proceedings - International Conference on Pattern Recognition*, pages 3269–3273, 2014.

[28] Feldman Naik, S., Doyle, S., Madabhushi, A., Tomaszewski, J. Automated nuclear and gland segmentation and gleason grading of prostate histology by integrating low, high-level and domain specific information. *In: Special Workshop on Computational Histopathology (CHIP), in conjunction with 5th IEEE Internat. Symp. on Biomedical Imaging (*, pages 1–8, 2008.

[29] Michaela Weingant, Hayley M. Reynolds, Annette Haworth, Catherine Mitchell, Scott Williams, and Matthew D. DiFranco. Ensemble prostate tumor classification in HE whole slide imaging via stain normalization and cell density estimation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9352(October):280–287, 2015.

[30] Patel Janakkumar Baldevbhai. Color Image Segmentation for Medical Images using L*a*b* Color Space. *IOSR Journal of Electronics and Communication Engineering*, 1(2):24–45, 2012.

[31] J. P. Gasparri, A. Bouchet, G. Abras, V. Ballarin, and J. I. Pastore. Medical image segmentation using the HSI color space and Fuzzy Mathematical Morphology. *Journal of Physics: Conference Series*, 332(1), 2011.

[32] Wurood A and Jbara and Rafah A Jaafar. MRI Medical Images Enhancement based on Histogram Equalization and Adaptive Histogram Equalization. *International Journal of Computer Trends and Technology*, 50(2):91–93, 2017.

[33] Gurleen Singh and Sukhpreet Kaur. Combination of Brightness Preserving Bi-Histogram Equalization and Discrete Wavelet Transform using LUV Color Space for Image Enhancement. *International Journal of Computer Applications*, 148(13):26–30, 2016.

[34] Saul McLeod. Likert Scale Likert Scale Examples How can you analyze data from a Likert Scale ? *Simply Psychology*, pages 1–3, 2008.

[35] William K Pratt. *Digital Image Processing: PIKS Inside*, volume 5. 2007.

[36] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, volume 8. Springer International Publishing, 2021.

[37] Guifang Lin and Wei Shen. Research on convolutional neural network based on improved Relu piecewise activation function. *Procedia Computer Science*, 131:977–984, 2018.

[38] Yu Han Liu. Feature Extraction and Image Recognition with Convolutional Neural Networks. *Journal of Physics: Conference Series*, 1087(6), 2018.

[39] Iqbal H. Sarker. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3):1–21, 2021.

[40] Foroogh Sharifzadeh, Gholamreza Akbarizadeh, and Yousef Seifi Kavian. Ship Classifi-

cation in SAR Images Using a New Hybrid CNN–MLP Classifier. *Journal of the Indian Society of Remote Sensing*, 47(4):551–562, 2019.

[41] Guangyu Jia, Hak Keung Lam, and Yujia Xu. Classification of COVID-19 chest X-Ray and CT images using a type of dynamic CNN modification method. *Computers in Biology and Medicine*, 134(April):104425, 2021.

[42] Dimpy Varshni, Kartik Thakral, Lucky Agarwal, Rahul Nijhawan, and Ankush Mittal. Pneumonia Detection Using CNN based Feature Extraction. *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*, 2019.

[43] Connor Shorten and Taghi M. Khoshgoftaar. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1), 2019.

[44] Łukasz Rączkowski, Marcin Możejko, Joanna Zambonelli, and Ewa Szczurek. ARA: accurate, reliable and active histopathological image classification framework with Bayesian deep learning. *Scientific Reports*, 9(1):1–12, 2019.

[45] Shereen Fouad, David Randell, Antony Galton, Hisham Mehanna, and Gabriel Landini. Epithelium and stroma identification in histopathological images using unsupervised and semi-supervised superpixel-based segmentation. *Journal of Imaging*, 3(4):1–18, 2017.

[46] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. 2017.

[47] Julián David Vargas López, Nicolás Toro-García, Juan Bernardo Gomez Mendoza, Paula Andrea Toro-Castaño, Rafael Pava-Marín, and Álex Enrique Pava-Ripoll. Histopathology color image processing in prostate carcinoma. page 16, 2020.

[48] Yash Sharmay, Lubaina Ehsany, Sana Syed, and Donald E. Brown. HistoTransfer: Understanding Transfer Learning for Histopathology. pages 1–4, 2021.

[49] Eirini Arvaniti, Kim S. Fricker, Michael Moret, Niels Rupp, Thomas Hermanns, Christian Fankhauser, Norbert Wey, Peter J. Wild, Jan H. Rüschoff, and Manfred Claassen. Automated Gleason grading of prostate cancer tissue microarrays via deep learning. *Scientific Reports*, 8(1):1–11, 2018.

[50] Yada Pruksachatkun, Jason Phang, Haokun Liu, Phu Mon Htut, Xiaoyi Zhang, Richard Yuanzhe Pang, Clara Vania, Katharina Kann, and Samuel R. Bowman. Intermediate-Task Transfer Learning with Pretrained Language Models: When and

Why Does It Work? pages 5231–5247, 2020.

[51] Zabit Hameed, Sofia Zahia, Begonya Garcia-Zapirain, José Javier Aguirre, and Ana María Vanegas. Breast cancer histopathology image classification using an ensemble of deep learning models. *Sensors (Switzerland)*, 20(16):1–17, 2020.

[52] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pages 1–14, 2015.

[53] Chiranjibi Sitaula and Sunil Aryal. Fusion of whole and part features for the classification of histopathological image of breast tissue. *Health Information Science and Systems*, 8(1):1–12, 2020.

[54] Wenchao Han, Carol Johnson, Mena Gaed, José A. Gómez, Madeleine Moussa, Joseph L. Chin, Stephen Pautler, Glenn S. Bauman, and Aaron D. Ward. Histologic tissue components provide major cues for machine learning-based prostate cancer detection and grading on prostatectomy specimens. *Scientific Reports*, 10(1):1–12, 2020.

[55] Shidan Wang, Alyssa Chen, Lin Yang, Ling Cai, Yang Xie, Junya Fujimoto, Adi Gazdar, and Guanghua Xiao. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. *Scientific Reports*, 8(1):1–9, 2018.

[56] Weiwei Shi, Yihong Gong, Xiaoyu Tao, Jinjun Wang, and Nanning Zheng. Improving CNN performance accuracies with min-max objective. *IEEE Transactions on Neural Networks and Learning Systems*, 29(7):2872–2885, 2018.

[57] Pedro Porto Buarque de Gusmão, Gianluca Francini, Skjalg Lepsøy, and Enrico Magli. Fast Training of Convolutional Neural Networks via Kernel Rescaling. pages 1–13, 2016.

[58] Yizhi Liu, Yao Wang, Ruofei Yu, Mu Li, Vin Sharma, and Yida Wang. Optimizing CNN model inference on CPUs. *Proceedings of the 2019 USENIX Annual Technical Conference, USENIX ATC 2019*, pages 1025–1039, 2019.

[59] Samuel S. Ogden and Tian Guo. Characterizing the Deep Neural Networks Inference Performance of Mobile Applications. 2019.

[60] Marc Macenko, Marc Niethammer, J. S. Marron, David Borland, John T. Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E. Thomas. A method for normalizing histology slides for quantitative analysis. *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, pages 1107–1110, 2009.