

Article

Impact of Image Preprocessing Methods and Deep Learning Models for Classifying Histopathological Breast Cancer Images

David Murcia-Gómez ¹, Ignacio Rojas-Valenzuela ¹  and Olga Valenzuela ^{2,*} ¹ School of Technology and Telecommunications Engineering, University of Granada, 18071 Granada, Spain² Department of Applied Mathematics, University of Granada, 18071 Granada, Spain

* Correspondence: olgavc@ugr.es

Abstract: Early diagnosis of cancer is very important as it significantly increases the chances of appropriate treatment and survival. To this end, Deep Learning models are increasingly used in the classification and segmentation of histopathological images, as they obtain high accuracy index and can help specialists. In most cases, images need to be preprocessed for these models to work correctly. In this paper, a comparative study of different preprocessing methods and deep learning models for a set of breast cancer images is presented. For this purpose, the statistical test ANOVA with data obtained from the performance of five different deep learning models is analyzed. An important conclusion from this test can be obtained; from the point of view of the accuracy of the system, the main repercussion is the deep learning models used, however, the filter used for the preprocessing of the image, has no statistical significance for the behavior of the system.

Keywords: deep learning; cancer; image preprocessing method; ANOVA



Citation: Murcia-Gómez, D.; Rojas-Valenzuela, I.; Valenzuela, O. Impact of Image Preprocessing Methods and Deep Learning Models for Classifying Histopathological Breast Cancer Images. *Appl. Sci.* **2022**, *12*, 11375. <https://doi.org/10.3390/app122211375>

Academic Editors: Fabio La Foresta and Jan Egger

Received: 30 September 2022

Accepted: 5 November 2022

Published: 9 November 2022

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



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

1. Introduction

In digital pathology, whole slide images (WSI) are considered the golden standard for primary diagnosis [1,2]. The use of computer-assisted image analysis tools is becoming mainstream for automatic quantitative or semi-quantitative analysis for pathologists [3–7].

Cancer is one of the critical public health issues around the world. According to the Global Burden of Disease (GBD) study, there have been more than 25 million cancer incidences and 10 million cancer deaths worldwide in 2020. These statistics indicate that cancer incidence expanded by 33% between 2007 and 2017 worldwide [8] and, for example, is projected to increase more than 92% in 2020 and 2040 in sub-Saharan Africa [9].

Breast cancer is one of the leading causes of cancer-related deaths worldwide and affects a large number of women today [10,11]. Invasive ductal carcinoma (IDC) is the most frequent subtype of all breast cancers [12]. As stated in [13], 934,870 new cancer cases and 287,270 cancer deaths in women are projected to occur in the United States in 2022, with an estimated 287,850 new cases of breast cancer (the leading cause, accounting for 31% of new cancer cases) and 43,250 deaths from breast cancer (the second leading estimated cause of cancer death, with breast cancer accounting for 15% in women). Incidence rates for breast cancer in women have slowly increased by about 0.5% per year since the mid-2000s [14].

Most women diagnosed with breast cancer are over the age of 50, but younger women can also develop breast cancer. About 1 in 8 women will be diagnosed with breast cancer during their lifetime. Survival rates for breast cancer are believed to have increased and the number of related deaths continues to decline, largely due to factors such as earlier detection, a new personalized treatment approach, and a better understanding of the disease [15,16].

Histopathology is a technique frequently used to diagnose tumors. It can identify the characteristics of how cancer looks under the microscope. Accurate detection and classification of some types of cancer is a critical task in medical imaging due to the

complexity of those tissues [17,18], concretely in breast cancer [19], which is the one that will be analyzed in this paper.

The diagnosis of breast cancer histopathology images with hematoxylin and eosin stained is non-trivial, labor-intensive, and often leads to a disagreement between pathologists [20,21]. Computer-assisted diagnosis systems contribute to helping pathologists improve diagnostic consistency and efficiency, especially in breast cancer using histopathological images [22–24].

The use of artificial intelligence tools, especially deep learning methods, to classify histopathological images of patients with breast cancer has been extensively and fruitfully analyzed in the literature [25–29]. In particular, convolutional neural networks (CNNs) are the most commonly used for histology image analysis tasks [30–33].

Once the database of the WSI images is available, a phase that is usually performed in different papers presented in the bibliography, is to carry out a pre-processing of said images. The main objective of this pre-processing step is to improve the behavior of the system (several alternatives for pre-processing can be used, among which it is worth mentioning to remove noise, blurring, enhance contrast, and, ultimately, improve the quality of the images to obtain better performance for the final deep learning algorithms). In Figure 1 a typical classification model pipeline is presented in which common techniques, such as data split, image pre-processing, tiling the whole slide images (WSI), and statistical evaluation with a test set, are highlighted.

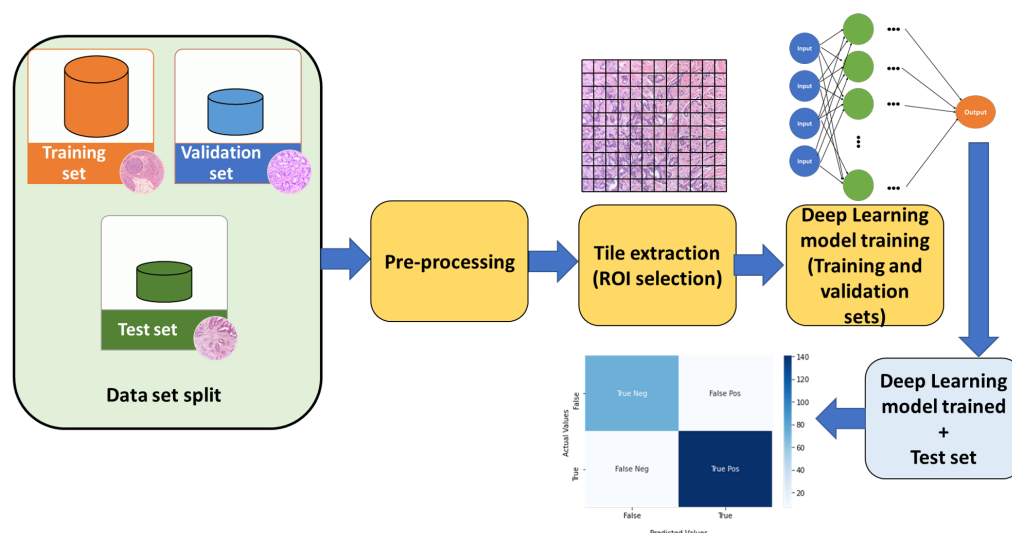


Figure 1. Illustrating typical framework for pathology WSI classification with deep learning model.

This paper proposes a comparative study of how different histopathological image preprocessing methods, and various deep learning models, affect the performance of the classification system in breast cancer. The methods used are divided into two groups: based on convolution filters and based on histogram equalization. These preprocessing are tested on a data set containing images of breast cancer.

Our aim with this article is to accurately determine the influence of the different deep learning models and filters (denoted as factors or ways of creating the deep learning system) for analyzing the influence on the behavior (accuracy) of the developed system (in terms of the AUC). To carry out this statistical analysis, the Analysis Of Variance tool (ANOVA), is used [34–37].

The rest of this paper is organized as follows. In Section 2, we present a short overview of some papers employing preprocessing methods in different contexts. Section 3 describes the dataset used and the preprocessing methods applied. In addition, the deep learning models used and the metrics used for their evaluation are presented. Section 4 explains how the experiments with the different models and filters were carried out. Finally, Section 5 provides a discussion of the results and concludes the paper.

2. Related Work

In problems with histopathologic images, it is increasingly common to use some form of preprocessing. In [38], the authors provide a detailed overview of graph-based deep learning for computational histopathology. They note that stain normalization is a common practice when using standard entity graph pathology workflows. Similarly, in [39] stain normalization is used as preprocessing for a classification model that predicts not only histological sub-types of endometrial carcinoma, but also molecular sub-types and 18 common genetic mutations based on histopathological images.

A tumor segmentation problem in lung cancer images is addressed in [40]. In the preprocessing stage, they employ a staining normalization to homogenize the images. Then, they apply another type of Gaussian filter as a color augmentation method to make the models used more robust.

A problem of binary classification of breast cancer images, obtaining an accuracy higher than 98%, is presented in [41]. The procedure the authors follow for its resolution consists of three stages: preprocessing, segmentation and classification. In the preprocessing stage, the authors use a fuzzy equalization of the histogram of the images to overcome unwanted over-enhancement and noise amplification by preserving local information of the original image.

The segmentation of tumors in Cervical Cancer images was analyzed in [42]. To improve the quality of the images and thus increase the performance of the models, the development of an automatic framework to preprocess the images was proposed. This framework would include different preprocessing methods such as adaptive histogram equalization, Gaussian filters, or sharpen filters among others.

In [43], the authors propose a study in which they compare the results of applying different filters to retinal fundus images. The filters used are the following: Contrast Limited Adaptive Histogram Equalization (CLAHE), Adaptive Histogram Equalization (AHE), Median filter, Gaussian filter, Weiner filter, and Adaptive median filters. They also present the advantages and disadvantages of each method, giving a brief explanation of them. To quantify the efficiency of each filter they use the MSE and PSNR measures. Finally, they conclude that the best-performing filter is the adaptive median filter due to its high PSNR and low MSE.

Regarding the different models of deep learning, the alternatives that are frequently found in the literature and that have been successfully applied in histopathology images are VGG16, VGG19, ResNet50, MobilNet, DenseNet121, among others [44–46]. For example, in [47], the authors propose a cascaded deep learning framework for accurate nuclei segmentation in digitized microscopy images of histology slides, based on the VGG16 model and a soft-die loss function. In [48] a multi-class classification, using the 6B-Net model selected vector and ResNet50 selected vector was fused and applied to breast cancer. In [44], a review of the popular deep learning algorithms, including convolutional neural networks (CNNs), generative adversarial networks (GANs), and graph neural networks (GNNs) is presented, highlighting the applications in pathology.

3. Materials and Methods

This section is divided into four parts. In the first one, we will give a brief description of the dataset used in the experimentation. In the second, we will detail the preprocessing that has been applied, giving a short description of them. Then we will talk about the deep learning models chosen to carry out the experimentation, ending the section with the metrics that will be used to evaluate the performance of the models.

3.1. Dataset: Breast Histopathology Images

In this paper, the dataset of breast cancer images presented in [49] will be used. It consists of 277,524 color images of size 50×50 px. The images are annotated with a binary label indicating the presence of Invasive Ductal Carcinoma. In this case, a random undersampling balancing of the images has been performed as there are more than twice as

many healthy images as cancerous ones (198,738 healthy vs. 78,786 cancerous ones). After the balancing, the images of the train, validation, and test sets were also randomly selected, as shown in Table 1.

Table 1. Information about the number of breast cancer patches for each data split.

	Classes		
	Healthy (0)	Cancer (1)	Total
Train	63,050	63,050	126,100
Validation	7868	7868	15,736
Test	7868	7868	15,736
Discarded	119,952	0	119,952
Total	198,738	78,786	277,524

3.2. Preprocessing

The impact of the pre-processing techniques on deep learning frameworks is of great relevance [50]. There are a large number of proposals and methodologies in pre- and post-image processing in digital pathology [51]. In this section, we will present the most frequent ones that appear in the bibliography.

3.2.1. Convolution-Based

A convolution is a filter in which the neighborhood of each pixel is organized by weighted average, according to a matrix that we call the convolution kernel [52]. These kernels are variable in size and shape producing different effects on the image.

Mathematically, the convolution operation can be represented as follows [53].

Let $I = I(i, j)_{i=1, \dots, N}^{j=1, \dots, M}$ be a grayscale image. Let $M = M(k, l)_{k=1, \dots, p}^{l=1, \dots, q}$ be the convolution kernel matrix, where $p = 2n + 1, q = 2m + 1$ with n, m non negatives integers. The result of applying the convolution M to the image I is the image $I' = I'(i, j)_{i=1, \dots, N}^{j=1, \dots, M}$ where for each (x_0, y_0)

$$I'(x_0, y_0) = \sum_{a=-n}^n \sum_{b=-m}^m M(a, b)I(x_0 + a, y_0 + b)$$

where $1 \leq x_0 - n, x_0 + n \leq N, 1 \leq y_0 - m, y_0 + m \leq M$.

In this paper we have used some non-parametric filters from the Python Pillow library [54]. Although they are different in size and the effect they have on the image to which they are applied, they are all based on convolution and the only thing that varies is the kernel. Table 2 shows the matrices of the filters used, which are as follows:

1. **Blur.** Applies a convolution with a kernel of size 5×5 which will result in a slightly blurred image. Because of the shape of the kernel, the central pixels will be ignored, thus losing some information.
2. **Contour.** Convolute with a 3×3 filter. This filter detects the contours of shapes in images.
3. **Detail.** It applies a 3×3 size convolution, giving greater importance to the central pixel compared to its neighborhood, thus achieving an enhanced image.
4. **Edge Enhance.** Applies a convolution with a 3×3 kernel. It aims to improve the quality and definition of the edges of an image.
5. **Edge Enhance More.** As in the previous filter, a 3×3 convolution is applied to improve the definition of the edges. The only difference is that the central value is changed, making the edges better defined than with the previous filter.
6. **Sharpen.** This filter convolves a 3×3 kernel to generate a cleaner image. It is used to sharpen the edges to improve their quality. It even increases the contrast between the light and dark areas of the image to improve its characteristics.

Table 2. Convolution Kernel Matrices.

Blur	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$
Contour	$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$
Detail	$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 10 & -1 \\ 0 & -1 & 0 \end{pmatrix}$
Edge Enhance	$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 10 & -1 \\ -1 & -1 & -1 \end{pmatrix}$
Edge Enhance More	$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}$
Sharpen	$\begin{pmatrix} -2 & -2 & -2 \\ -2 & 32 & -2 \\ -2 & -2 & -2 \end{pmatrix}$

3.2.2. Histogram-Based

A histogram is a graphical representation that shows the frequency with which a certain intensity appears in an image, i.e., it represents the number of pixels of each intensity in an image. On the X-axis is usually placed the intensity of the pixels while on the Y-axis the frequency with which the pixels appear. The intensity level usually ranges from 0 to 255. For a grayscale image, there is only one histogram, while an RGB color image will have three histograms, one for each color.

The methods in this section are based on applying a transformation on the image, so that the histograms acquire a uniform distribution. This type of transformation is used to improve the contrast of the images [55].

1. **Histogram Equalization (HE).** Histogram equalization is one of the most common ways to enhance the contrast of an image. This is because it is easy to implement and does not require too many computational resources. It is performed by mapping the gray levels of the image based on the probability distribution of the input gray levels. The detailed procedure can be found at [56]. However, one of the main problems it presents is that it tends to change the brightness of the image significantly and therefore causes the output image to be saturated with very bright or dark intensity values.
2. **Contrast Limited Histogram Equalization (CLAHE).** Conventional histogram equalization considers the overall contrast of the image. In many cases, it is not a good idea because when the histogram presents a distribution that is concentrated in a range, it does not work properly [57]. To solve this problem, adaptive histogram equalization is used, which is carried out as follows. First, the image is divided into small blocks of a predefined size. Then, each of these blocks is histogram equalized as in the previous section. So in an area, the histogram will be limited to a small region unless there is noise in which case it will be amplified. To avoid this behavior of the algorithm, contrast limiting is applied. If any bin in the histogram is above a certain contrast threshold, those pixels are clipped and evenly distributed to other bins before equalization is applied. Then, to remove edge artifacts, bilinear interpolation is applied. The default parameters offered by the OpenCV library for this method have been used, namely: clipLimit = 2.0, tileGridSize = (8,8).

- Fuzzy Histogram Equalization (FHE).** To perform contrast enhancement of the images by means of fuzzy histogram equalization, the gray level intensities of the image are mapped to a fuzzy space using membership functions to enhance the contrast [58]. Then the fuzzy plane is mapped to the gray intensities of the image. This process is intended to obtain an image with better contrast than the original image by giving more importance to the gray levels that are closer to the mean than to those that are farther away. Fuzzy image processing consists of three stages: fuzzification, membership plane operations, and defuzzification. The first stage consists of assigning certain membership values to the images based on some attributes of the images such as brightness, homogeneity, etc. Then, using a fuzzy approach, the membership values assigned in the first stage are modified. Finally, in the defuzzification stage, the membership values that have been treated in the previous stage are decoded and transformed back into the gray level plane. In general, dark pixels are assigned low membership values and bright pixels are assigned high membership values [59]. The pipeline used is from [60].

An example of the effect of the different filters, both in healthy images (category 0) and in images with cancer (category 1), for the different filters analyzed, is presented in Figure 2.

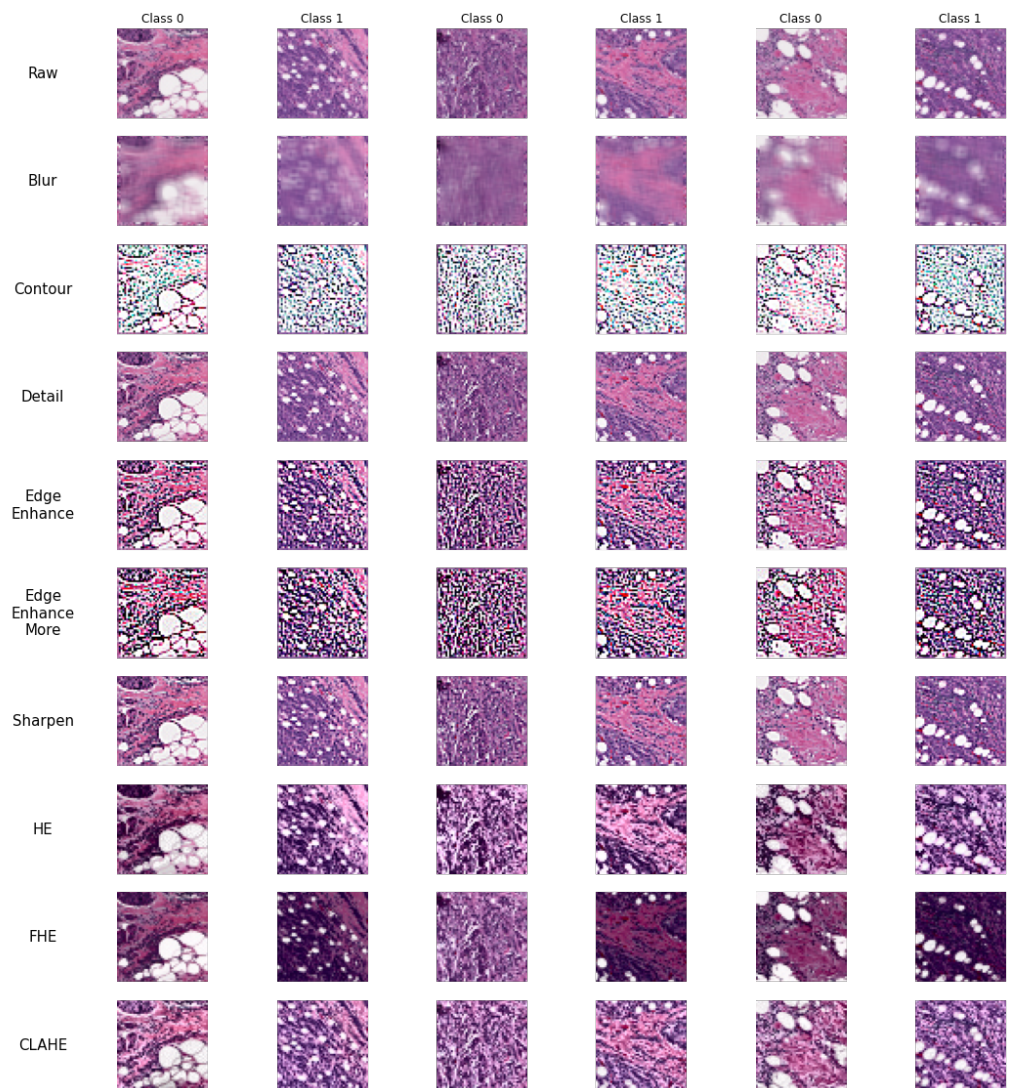


Figure 2. Some examples of how filters act on different images.

3.3. Models Used

Five deep learning models have been selected to use architectures that are diverse in size and performance. We have not tried to make an exhaustive analysis of which models to select, but rather, we have taken models that normally show good performance in image classification tasks and whose architectures are varied in the number of parameters and layers. Thus, the models used were *VGG16* and *VGG19* [61], *MobileNet* [62], *ResNet50* [63] and *DenseNet121* [64]. These five models have been widely used in the bibliography [65–71].

1. **VGG16 and VGG19.** Karen Simonyan and Andrew Zisserman proposed the idea of the VGG network in 2013 and presented the current model based on this idea at the 2014 ImageNet Challenge [61]. VGG is a special convolutional network designed for classification. They named it VGG after the department of the Visual Geometry Group at Oxford College to which they belonged, and because it is a convolutional neural network that is either 16 layers deep (VGG-16) or 19 layers deep (VGG-19). The VGG16 and VGG19 architecture in Deep Learning is a rather simple architecture that uses only blocks consisting of an increasing number of convolutional layers with filters of size 3×3 . Moreover, to reduce the size of the activation maps obtained, max-pooling blocks are inserted between the convolutional layers, reducing the size of these activation maps by half. Finally, a classification block is used, consisting of two dense layers of 4096 neurons each and a final layer, the output layer, of 1000 neurons.
2. **MobileNet.** This model was proposed in 2017 and it is based on a simplified architecture with depth-separable convolutions. The model was originally developed for mobile imaging applications. The architecture of MobileNet initially includes a convolutional layer with 32 filters, followed by 19 layers. MobileNet uses 3×3 filters, dropout (an operation in which neurons generate partial learning of the network to avoid overtraining), and normalization during training [72,73]
3. **ResNet.** This deep learning model is a deep convolutional neural network with 50 layers. The architecture of ResNet-50 starts with a convolution with a kernel size of 7×7 and 64 different kernels, all with a step size of 2, so we get one layer. Then there are different convolutional layers with different kernels [74,75].
4. **DenseNet-121.** The denseNet-121 model of densely connected convolutional networks, proposed in late 2017, consists of networks with blocks where each block has direct access to the gradient of the cost function, the function that provides the error, and the initial input of the block. Of course, there are also convolutional layers and pooling layers [76].

Table 3 summarizes different deep learning models that were used in this paper alongside pre-trained weights (using information from Keras Applications). Depth refers to the topological depth of the network (including activation layers, batch normalization layers, etc).

Table 3. Summary of Different Neural Network Architectures.

Model	Size (MB)	Parameters	Depth	Main Features
VGG16	528	138.4M	16	VGG16 and VGG19 use multiple convolutional layers with a small kernel size (3×3) to replace a single convolutional layer with a large kernel size. These convolutional layers are accompanied by max-pooling layers, each 2×2 in size, to reduce the size of the filters during training.
VGG19	549	143.7M	19	Although both architectures are very similar and follow the same logic, VGG19 has a higher number of convolution layers.
ResNet50	98	25.6M	107	It provides an innovative way to add more convolutional layers to a CNN, without falling into the vanishing gradient problem, by employing the residual block concept.
MobileNet	16	4.3M	55	It is a simple, yet efficient and computationally inexpensive convolutional neural network for mobile vision applications.
DenseNet121	33	8.1M	242	DenseNets simplify the connectivity pattern between layers. Each layer is connected directly with every other layer.

3.4. Experimental Settings

A total of three different runs were performed, using a transfer learning technique, for each filter and each model by randomly selecting the training, validation, and test sets.

Different scenarios can be considered for transfer learning. The first scenario would be to have a large dataset (in our case images), but different from the pre-trained model's dataset and to be able to train the entire model, being useful to initialize the model from a pre-trained model, using its architecture and weights. The second scenario would be when the set of images that you want to train has a certain similarity with the images that were used to train the model, or when you have a small set of images to train the new system. In this case, it is usually chosen to train some layers and leave others frozen. Finally, the third scenario would be to have a small set of images, similar to the pre-trained model's dataset. In that case, they usually freeze the convolutional base.

The scenario used in this article has been the first scenario: the model is initialized from a pre-trained model (using its architecture and weights), and the entire model is trained/adjusted (no layer was frozen).

Therefore, the procedure that was followed for each filter and deep learning model is the following. First, the model pre-trained with the Imagenet dataset is loaded, removing the last default layers, and replacing them with layers that offer an adequate solution to the problem to be solved (flatten layer). On the other hand, after some experiments in cases where the models suffered from overfitting, measures would be incorporated to avoid its behavior, so we embed regularization kernels and a dropout layer that 20% of the neurons were removed. Finally, a fully connected layer whose activation function is a sigmoid. This results in a single output. Additionally, in this last layer, l_1 and l_2 kernels have been applied to regulate training. These kernels are penalty factors added to the loss function so that abrupt changes in the values of the weights are penalized during training to avoid over-fitting the models. Therefore, the usual transfer learning procedure was used. Relying on prior learning, it avoids starting from scratch and allows us to build accurate models in a time-saving manner [77–79].

The input to the different deep learning models is the breast cancer histopathology images. These images received the appropriate scaling for the expected input to each neural network. An Adam optimizer with a learning rate of 10^{-5} was used. During training, the binary cross entropy was used as the loss function and a total of 50 epochs were iterated for each model. However, in general, not so many iterations were performed because early stopping with a latency of 15 was used in terms of validation accuracy, with the weights of the best result recovered at the end of training. A batch size of 128 was used.

3.5. Evaluation Metrics

To measure the performance of the models and compare the effect of preprocessing on them, we will use the metrics accuracy, precision, recall, and AUC.

Accuracy. It is a general measure that quantifies the proportion of hits in the classification. Given the number of true positives (*TP*), false positives (*FP*), true negatives (*TN*), and false negatives (*FN*), the accuracy is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision. Quantifies actual positives versus all images that have been classified as positive.

$$Precision = \frac{TP}{TP + FP}$$

Recall. It measures the number of well-classified positives versus actual positives.

$$Recall = \frac{TP}{TP + FN}$$

AUC. In addition, we will use the area under the curve. Both the ROC curve and the *AUC* are a measure of the overall performance of a binary classification model. It measures the ability of the model to correctly classify positives at different thresholds. The higher the *AUC* the better the model is at predicting. A perfect model would have an *AUC* of 1 and a random classifier would have a value of 0.5 [80].

4. Experimental Results

Table 4 shows all the information provided by the runs performed. In addition to the Accuracy and AUC metrics considered in the analysis of ANOVA, the complementary measures for Precision and Recall are also provided.

To perform the statistical study, a selection is made from a set of alternatives representative of each of the factors considered for performing the deep learning training procedure. The factors considered in the analysis are listed in Table 5. By analyzing the different levels of each of these factors, it is possible to determine their influence on the performance of the analyzed system (accuracy) when different alternatives for preprocessing the images and also different deep learning models are presented.

Table 4. Comparative analysis ($\mu \pm \sigma$) of the performance of the different models and filters employed. In bold, the best result of all models for a fixed filter.

Metrics	Models	Raw	Blur	CLAHE	Contour	Detail	Edge Enhance	Edge Enhance More	FHE	HE	Sharpen
Accuracy	VGG16	86.93 ± 0.17	85.69 ± 0.23	86.22 ± 0.35	85.62 ± 0.30	86.85 ± 0.22	86.49 ± 0.35	86.22 ± 0.45	86.48 ± 0.37	85.81 ± 0.16	86.85 ± 0.27
	VGG19	86.84 ± 0.15	85.73 ± 0.12	86.24 ± 0.25	85.80 ± 0.33	86.55 ± 0.28	86.38 ± 0.30	86.25 ± 0.31	86.39 ± 0.17	85.63 ± 0.34	86.71 ± 0.23
	ResNet50	86.98 ± 0.16	86.00 ± 0.05	85.89 ± 0.17	85.39 ± 0.33	86.85 ± 0.10	86.31 ± 0.22	85.74 ± 0.04	86.39 ± 0.07	85.25 ± 0.09	86.95 ± 0.14
	MobilNet	86.44 ± 0.08	85.53 ± 0.58	85.43 ± 0.25	84.37 ± 0.27	86.16 ± 0.15	86.09 ± 0.40	85.55 ± 0.05	85.84 ± 0.17	84.93 ± 0.09	86.38 ± 0.08
	DenseNet121	85.62 ± 0.12	84.69 ± 0.19	84.66 ± 0.18	83.88 ± 0.20	85.35 ± 0.36	84.89 ± 0.13	84.43 ± 0.01	85.11 ± 0.12	84.23 ± 0.23	85.38 ± 0.20
Precision	VGG16	87.39 ± 0.42	86.34 ± 0.82	86.89 ± 1.23	85.20 ± 0.50	86.23 ± 1.23	85.66 ± 1.62	85.80 ± 1.45	86.34 ± 0.46	86.35 ± 1.21	86.19 ± 0.98
	VGG19	86.24 ± 1.57	86.35 ± 0.57	84.94 ± 1.79	85.78 ± 0.40	84.54 ± 0.97	86.50 ± 0.40	86.35 ± 0.71	85.41 ± 1.20	83.98 ± 1.36	85.35 ± 1.05
	ResNet50	86.21 ± 0.46	85.67 ± 0.56	85.43 ± 0.17	84.32 ± 0.65	86.42 ± 0.23	85.84 ± 0.34	85.12 ± 0.17	86.03 ± 0.48	84.99 ± 0.16	86.37 ± 0.23
	MobilNet	86.51 ± 0.40	84.24 ± 1.68	84.17 ± 2.05	83.19 ± 1.91	85.04 ± 1.89	86.09 ± 0.12	84.26 ± 0.19	84.79 ± 1.28	84.87 ± 1.32	84.56 ± 0.72
	DenseNet121	85.24 ± 0.35	84.25 ± 0.33	83.85 ± 0.52	83.20 ± 0.30	85.16 ± 0.42	84.43 ± 0.10	83.90 ± 0.15	84.85 ± 0.10	83.68 ± 0.16	85.20 ± 0.35
Recall	VGG16	86.32 ± 0.38	84.82 ± 0.92	85.37 ± 2.22	86.23 ± 1.02	87.76 ± 1.21	87.75 ± 1.50	86.86 ± 1.33	86.68 ± 0.97	85.13 ± 1.84	87.83 ± 0.90
	VGG19	87.76 ± 2.52	84.90 ± 1.00	88.21 ± 2.28	85.85 ± 0.91	89.49 ± 0.78	86.22 ± 0.98	86.13 ± 0.64	87.81 ± 1.43	88.14 ± 1.63	88.71 ± 1.06
	ResNet50	88.04 ± 0.35	86.48 ± 0.77	86.54 ± 0.47	86.96 ± 0.20	87.44 ± 0.43	86.96 ± 0.08	86.61 ± 0.28	86.90 ± 0.81	85.61 ± 0.15	87.78 ± 0.30
	MobilNet	86.34 ± 0.36	87.50 ± 1.20	87.44 ± 2.48	86.30 ± 2.65	87.89 ± 2.63	85.87 ± 0.53	87.43 ± 0.38	87.43 ± 2.18	85.08 ± 2.08	89.06 ± 0.98
	DenseNet121	86.43 ± 0.64	85.35 ± 0.32	85.86 ± 0.43	84.91 ± 0.12	85.62 ± 0.59	85.55 ± 0.39	85.21 ± 0.22	85.48 ± 0.29	85.05 ± 0.43	85.66 ± 0.20
AUC	VGG16	94.15 ± 0.14	93.24 ± 0.16	93.63 ± 0.10	92.75 ± 0.26	93.93 ± 0.26	93.66 ± 0.22	93.32 ± 0.23	93.60 ± 0.32	93.12 ± 0.06	93.86 ± 0.32
	VGG19	94.13 ± 0.06	94.29 ± 1.63	93.63 ± 0.04	92.88 ± 0.39	93.96 ± 0.18	93.61 ± 0.27	93.34 ± 0.37	93.77 ± 0.04	93.07 ± 0.19	94.02 ± 0.12
	ResNet50	93.93 ± 0.15	93.29 ± 0.08	93.98 ± 1.64	92.47 ± 0.35	93.87 ± 0.22	93.53 ± 0.07	92.93 ± 0.04	93.41 ± 0.21	92.48 ± 0.16	93.86 ± 0.03
	MobilNet	93.35 ± 0.16	92.57 ± 0.63	92.38 ± 0.38	91.57 ± 0.47	93.32 ± 0.38	92.77 ± 0.15	92.41 ± 0.01	92.91 ± 0.15	92.09 ± 0.12	93.50 ± 0.07
	DenseNet121	92.92 ± 0.09	92.16 ± 0.08	92.24 ± 0.07	91.22 ± 0.22	92.60 ± 0.18	92.33 ± 0.07	91.89 ± 0.07	92.51 ± 0.03	91.87 ± 0.12	92.58 ± 0.16

Table 5. Variables used in the statistical analysis. All the possible configurations of factors levels.

Factors	Filter	Model
Levels of the factors	Raw	VGG16
	Blur	VGG19
	CLAHE	ResNet50
	Contour	MobileNet
	Detail	DenseNet121
	Edge Enhance	
	Edge Enhance More	
	FHE	
	HE	
	Sharpen	

Table 6 gives the two-way variance analysis for the whole set of processing examples of the deep learning system analyzed in this contribution. The ANOVA table containing the sum of squares, degrees of freedom, mean square, test statistics, etc., represents the initial analysis in a compact form. This kind of tabular representation is customarily used to set out the results of the ANOVA calculations. The table ANOVA decomposes the variability of the AUC into contributions due to different factors. Because the sums of squares are of type III (the default), the contribution of each factor is measured after the effects of all other factors have been removed. The P-values test the statistical significance of each factor. Since a P-value is less than 0.05, this factor has a statistically significant effect on the AUC at a 95% confidence level.

Table 6. Analysis of Variance for the dependent variable that analyzes the AUC of the system when the model and the filter are modified.

Source	Sum of Squares	Df	Mean Square	F-Ratio	p-Value
Main Effects					
A: Filter	29.2101	9	3.24557	1.52	0.1460
B: Model	95.588	4	23.897	11.21	0.0000
Residual	289.987	136	2.13226		
Total (Corrected)	414.785	149			

Of all the information presented in the table ANOVA, the researcher's main interest is likely to be focused on the value in the "F-ratio" and "p-value" columns. If the numbers in this column are smaller than the critical value set by the experimenter, the effect is considered significant. This p-value is often set at 0.05. Any value lower than this leads to significant effects, while any value higher than this leads to non-significant effects [35]. When effects are found to be significant using the procedure described above, it means that there is more difference between the means than would be expected from chance alone. With respect to the experiment above, this would mean that the preprocessing filter of the WSI images has no significant effect on the accuracy of the system. As can be seen from Table 6, the model factor is statistically significant.

Thus, a detailed analysis will be performed now for each of the factors examined, using the Multiple Range Test for the two factors (model and filter) in which a multiple comparison procedure is carried out to determine which means are significantly different from which others. The method currently being used to discriminate among the means is Fisher's least significant difference (LSD) procedure [36]. A summary of this procedure can be seen in Table 7.

Table 7. Multiple Range Tests for the different factors. Method: 95.0 percent LSD.

Factors				
Filter	Count	LS means	Ls sigma	Homogeneous Groups
Contour	15	92.1773	0.377028	X
Sharpen	15	92.2287	0.377028	X X
HE	15	92.5227	0.377028	X X X
Edge Enhance More	15	92.776	0.377028	X X X
Blur	15	93.1107	0.377028	X X X
CLAHE	15	93.1713	0.377028	X X X
Edge Enhance	15	93.1793	0.377028	X X X
Raw	15	93.2287	0.377028	X X X
FHE	15	93.238	0.377028	X X
Detail	15	93.5333	0.377028	X X
Model	Count	LS means	Ls sigma	Homogeneous Groups
DenseNet121	30	91.5653	0.266599	X
MobileNet	30	92.451	0.266599	X
ResNet50	30	93.374	0.266599	X
VGG16	30	93.5247	0.266599	X
VGG19	30	93.668	0.266599	X

From this test can be observed that there exist three homogenous groups (identified using columns of X's, with intersection) for the variable filter. All of the groups have intersections (therefore, different levels of these groups have the same behavior over the AUC variable). These groups are:

- Contour, Sharpen, HE, Edge Enhance More, Blur, CLAHE, Edge Enhance, Raw
- Sharpen, HE, Edge Enhance More, BLUR, CLAHE, EDGE ENHANCE, raw, FHE
- HE, Edge Enhance More, BLUR, CLAHE, Edge Enhance, Raw, FHE, Detail

The levels of these 3 groups are not significant from a statistical point of view (analyzing the behavior of the dependent output variable) having an intersection between all of them. This information can be graphically represented in Figure 3.

On the other hand, with respect to the Model factor, there are three different groups without intersection.

- DenseNet121
- MobileNet
- ResNet50, VGG16, VGG19

The best group (best AUC) is the third. VGG19 obtains the best results. This factor is statistically significant. Figure 4 shows that, unlike the filter factor, the groups are well differentiated.

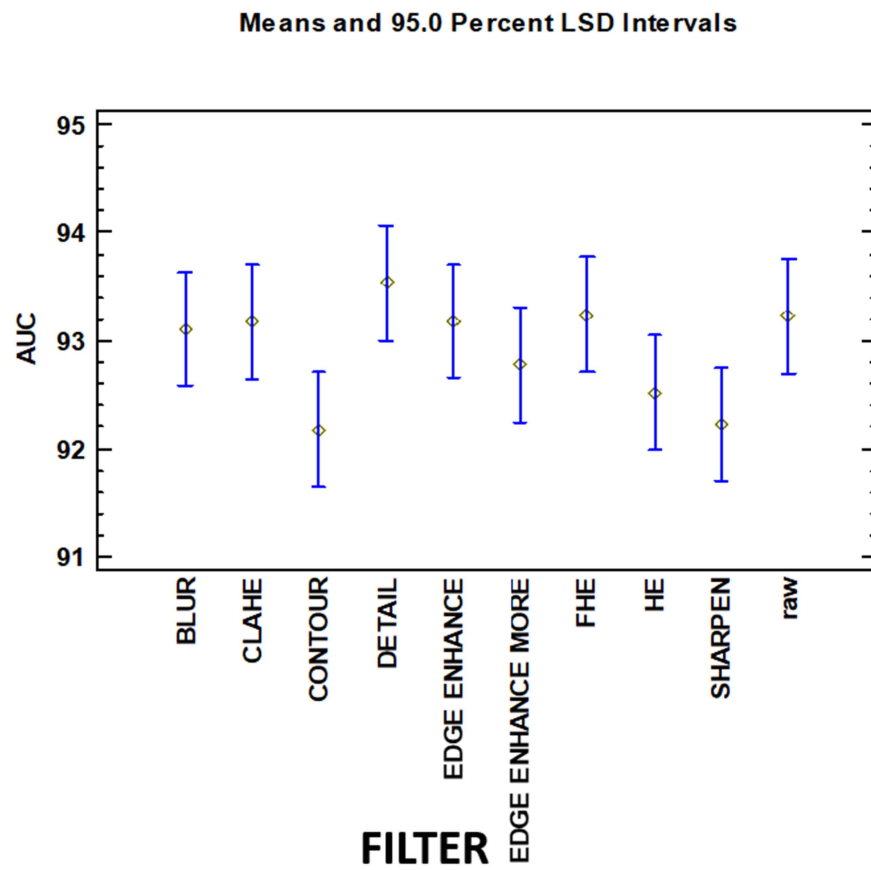


Figure 3. Multiple Range Tests for variable Filter.

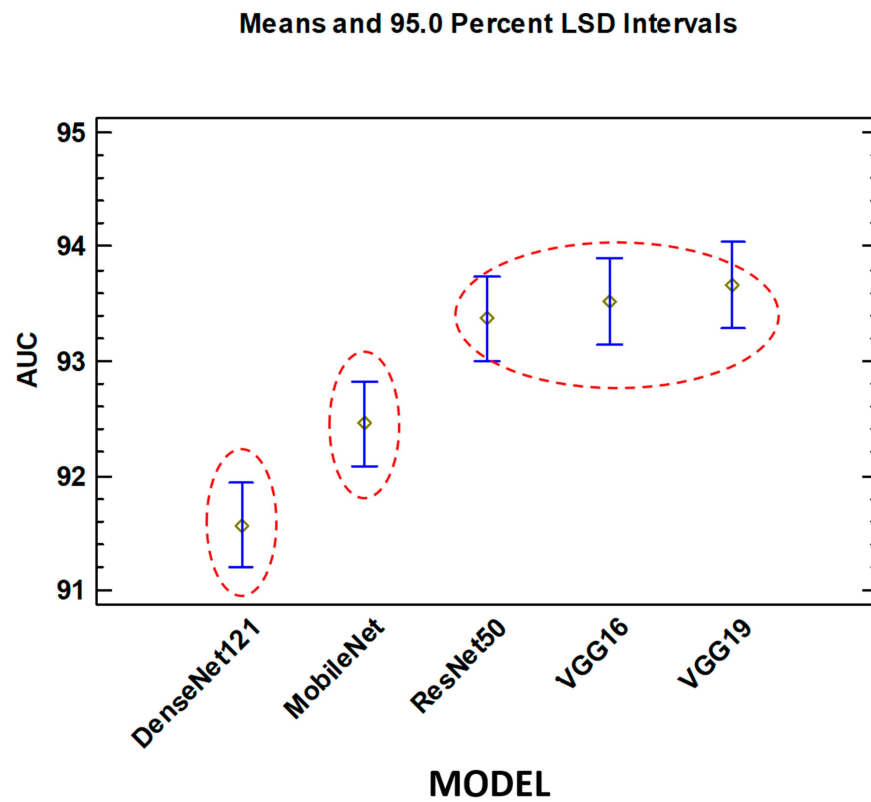


Figure 4. Multiple Range Tests for variable Model.

5. Discussion

As mentioned in Section 2, there are a large number of publications in the literature on the use of deep learning systems in medical image classification, especially in histopathology. Image pre-processing is an important step, and as can be seen from the bibliography, there are a large number of alternatives for its implementation (in this paper we have used the most frequently presented ones in the literature). On the other hand, there are a large number of deep learning models in the literature, and to perform statistical analysis with great rigor and accuracy, all possible combinations of filters and models must be simulated (in several repetitions), the behavior analyzed, and the error rates measured. The computational cost and time required for each simulation are high. For this reason, it was decided to select these five models that have been widely used in the literature. However, to the best of our knowledge, there is no exhaustive statistical analysis in the literature that attempts to analyze what relevance or impact it has on the behavior of the system to use the different alternatives of deep learning models and preprocessing of histological images for the problem of breast cancer.

As mentioned in Section 4, the analysis of the p -value in Table 6 yields a value of 0.146 for the filter technique and a value of 0.0000 for the model of deep learning. This means that the choice of the different deep learning model alternatives has a statistically significant influence on the behavior of the system (measured by the AUC index). On the other hand, for the filter effect, it can be stated that it is not statistically significant, and therefore, the different methods have similar behavior to obtain the AUC value. This means that the designer of a deep learning system, for the problem of histopathological cancer images studied, must focus more attention on the deep learning model to be used, than on the processing systems used. The three homogeneous groups of filter types intersect each other, therefore, from a statistical point of view they are equivalent in terms of their behavior on the AUC output variable. It is also important to note that different convolution-based and histogram-based filter alternatives have been used. These alternatives, of these two major types of pre-processing, are mixed in the three homogeneous groups of Table 7. No statistically significant difference was found when using convolution-based or histogram-based types, both methods have similar behavior for the output variable.

For the deep learning model type factor, there are also three groups, but, in this case, there are statistically significant differences between them. For this particular problem, and when analyzing and discussing the results, it can be seen that the group that achieves the best results (consisting of VGG16, VGG19, and ResNet50) are deep learning models that have a large number of parameters and therefore their size (measured in MB) and complexity is high. This does not carry over to the depth parameter (depth refers to the topological depth of the network, including activation layers, batch normalization layers, etc).

As might be initially expected, VGG16 and VGG19 behave similarly from a statistical point of view, and both produce the best results. While it is true that both have a larger number of parameters out of the five deep learning models analyzed, their depth topology is still the lowest. As strengths of this paper, we must highlight the novelty and robustness of performing a comprehensive statistical analysis of the impact that the different pre-processing algorithms and deep learning models have on the classification of histopathological images in breast cancer. As a possible weakness of the study, we point out that it would have been interesting to analyze other models of deep learning systems, preprocessing, and even other pathologies. This weakness is due to the high computational cost and time required to run multiple simulations for each of the combinations of filters and deep learning models.

6. Conclusions and Future Works

There is an extensive bibliography on the use of artificially intelligent systems for the classification and decision support of histopathological images. In particular, various deep learning models and image preprocessing algorithms have been proposed and used to obtain accurate systems. However, it is necessary to perform a comprehensive statistical

analysis of the influence of the preprocessing stage and deep learning models on the system behavior.

In this paper, it was proposed to use 50 different combinations (5 deep learning models and 10 image preprocessing alternatives) to statistically analyze the behavior of the system accurately through the ANOVA analysis. Accuracy, precision, recall, and AUC metrics were used to quantify the performance of the models. Using the ANOVA test, it was shown that the different alternatives of the filter have similar behavior in the accuracy indexes of the system. Furthermore, with the ANOVA test, it became clear that the choice of the architecture of the deep learning models used is statistically relevant. In future work, we propose to perform a similar statistical analysis and examine a wider variety of architectures to determine which models perform best.

Author Contributions: Conceptualization, D.M.-G., I.R.-V. and O.V.; methodology, D.M.-G., I.R.-V. and O.V.; writing—original draft preparation, D.M.-G., I.R.-V. and O.V.; writing—review and editing, D.M.-G., I.R.-V. and O.V.; visualization, D.M.-G.; supervision, O.V.; project administration, O.V.; funding acquisition, O.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Spanish Ministry of Sciences, Innovation, and Universities under Project PID2021-128317OB-I00 in collaboration with the Government of Andalusia under Project P20-00163.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available online in Breast Cancer Dataset. Details in Section 3.1. Dataset: Breast Histopathology Images.

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

References

1. Litjens, G.; Sánchez, C.I.; Timofeeva, N.; Hermsen, M.; Nagtegaal, I.; Kovacs, I.; van de Kaa, C.H.; Bult, P.; van Ginneken, B.; van der Laak, J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci. Rep.* **2016**, *6*, 26286. [[CrossRef](#)] [[PubMed](#)]
2. Komura, D.; Ishikawa, S. Machine Learning Methods for Histopathological Image Analysis. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 34–42. [[CrossRef](#)] [[PubMed](#)]
3. Lerousseau, M.; Vakalopoulou, M.; Classe, M.; Adam, J.; Battistella, E.; Carré, A.; Estienne, T.; Henry, T.; Deutsch, E.; Paragios, N. Weakly supervised multiple instance learning histopathological tumor segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2020; pp. 470–479.
4. Gurcan, M.; Boucheron, L.; Can, A.; Madabhushi, A.; Rajpoot, N.; Yener, B. Histopathological Image Analysis: A Review. *IEEE Rev. Biomed. Eng.* **2009**, *2*, 147–171. [[CrossRef](#)] [[PubMed](#)]
5. Madabhushi, A.; Lee, G. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Med. Image Anal.* **2016**, *33*, 170–175. [[CrossRef](#)]
6. Veta, M.; Pluim, J.P.W.; van Diest, P.J.; Viergever, M.A. Breast Cancer Histopathology Image Analysis: A Review. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 1400–1411. [[CrossRef](#)]
7. Mobark, N.; Hamad, S.; Rida, S.Z. CoroNet: Deep Neural Network-Based End-to-End Training for Breast Cancer Diagnosis. *Appl. Sci.* **2022**, *12*, 7080. [[CrossRef](#)]
8. Fitzmaurice, C.E. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: A systematic analysis for the global burden of disease study. *JAMA Oncol.* **2019**, *5*, 1749–1768.
9. World-Health-Organization. Estimated number of new cases from 2020 to 2040, Incidence, Both sexes, age [0–85+]. *Int. Agency Res. Cancer* **2022**, *18*, 4473.
10. Zhang, C.; Bai, Y.; Yang, C.; Cheng, R.; Tan, X.; Zhang, W.; Zhang, G. Histopathological image recognition of breast cancer based on three-channel reconstructed color slice feature fusion. *Biochem. Biophys. Res. Commun.* **2022**, *619*, 159–165. [[CrossRef](#)]
11. Budak, Ü.; Cömert, Z.; Rashid, Z.N.; Şengür, A.; Çıbuk, M. Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images. *Appl. Soft Comput.* **2019**, *85*, 105765. [[CrossRef](#)]
12. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
13. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **2022**, *72*, 7–33. [[CrossRef](#)] [[PubMed](#)]

14. Pfeiffer, R.M.; Webb-Vargas, Y.; Wheeler, W.; Gail, M.H. Proportion of U.S. Trends in Breast Cancer Incidence Attributable to Long-term Changes in Risk Factor Distributions. *Cancer Epidemiol. Biomark. Prev.* **2018**, *27*, 1214–1222. [[CrossRef](#)] [[PubMed](#)]
15. Fisher, B.; Anderson, S.; Bryant, J.; Margolese, R.G.; Deutsch, M.; Fisher, E.R.; Jeong, J.H.; Wolmark, N. Twenty-Year Follow-up of a Randomized Trial Comparing Total Mastectomy, Lumpectomy, and Lumpectomy plus Irradiation for the Treatment of Invasive Breast Cancer. *N. Engl. J. Med.* **2002**, *347*, 1233–1241. [[CrossRef](#)]
16. Cristofanilli, M.; Budd, G.T.; Ellis, M.J.; Stopeck, A.; Matera, J.; Miller, M.C.; Reuben, J.M.; Doyle, G.V.; Allard, W.J.; Terstappen, L.W.; et al. Circulating Tumor Cells, Disease Progression, and Survival in Metastatic Breast Cancer. *N. Engl. J. Med.* **2004**, *351*, 781–791. [[CrossRef](#)]
17. Guleria, S.; Shah, T.U.; Pulido, J.V.; Fasullo, M.; Ehsan, L.; Lippman, R.; Sali, R.; Mutha, P.; Cheng, L.; Brown, D.E.; et al. Deep learning systems detect dysplasia with human-like accuracy using histopathology and probe-based confocal laser endomicroscopy. *Sci. Rep.* **2021**, *11*, 5086. [[CrossRef](#)]
18. Bansal, K.; Bathla, R.K.; Kumar, Y. Deep transfer learning techniques with hybrid optimization in early prediction and diagnosis of different types of oral cancer. *Soft Comput.* **2022**, *26*, 11153–11184. [[CrossRef](#)]
19. Chen, C.; Zheng, S.; Guo, L.; Yang, X.; Song, Y.; Li, Z.; Zhu, Y.; Liu, X.; Li, Q.; Zhang, H.; et al. Identification of misdiagnosis by deep neural networks on a histopathologic review of breast cancer lymph node metastases. *Sci. Rep.* **2022**, *12*, 13482. [[CrossRef](#)]
20. Sirinukunwattana, K.; Raza, S.E.A.; Tsang, Y.W.; Snead, D.R.J.; Cree, I.A.; Rajpoot, N.M. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. *IEEE Trans. Med Imaging* **2016**, *35*, 1196–1206. [[CrossRef](#)]
21. Munien, C.; Viriri, S. Classification of Hematoxylin and Eosin-Stained Breast Cancer Histology Microscopy Images Using Transfer Learning with EfficientNets. *Comput. Intell. Neurosci.* **2021**, *2021*, 5580914. [[CrossRef](#)]
22. Araújo, T.; Aresta, G.; Castro, E.; Rouco, J.; Aguiar, P.; Eloy, C.; Polónia, A.; Campilho, A. Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS ONE* **2017**, *12*, e0177544. [[CrossRef](#)] [[PubMed](#)]
23. Becker, A.S.; Marcon, M.; Ghafoor, S.; Wurnig, M.C.; Frauenfelder, T.; Boss, A. Deep Learning in Mammography. *Investig. Radiol.* **2017**, *52*, 434–440. [[CrossRef](#)] [[PubMed](#)]
24. Nassif, A.B.; Talib, M.A.; Nasir, Q.; Afadar, Y.; Elgendy, O. Breast cancer detection using artificial intelligence techniques: A systematic literature review. *Artif. Intell. Med.* **2022**, *127*, 102276. [[CrossRef](#)]
25. Han, Z.; Wei, B.; Zheng, Y.; Yin, Y.; Li, K.; Li, S. Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model. *Sci. Rep.* **2017**, *7*, 4172. [[CrossRef](#)] [[PubMed](#)]
26. Sudharshan, P.; Petitjean, C.; Spanhol, F.; Oliveira, L.E.; Heutte, L.; Honeine, P. Multiple instance learning for histopathological breast cancer image classification. *Expert Syst. Appl.* **2019**, *117*, 103–111. [[CrossRef](#)]
27. Hamidinekoo, A.; Denton, E.; Rampun, A.; Honnor, K.; Zwiggelaar, R. Deep learning in mammography and breast histology, an overview and future trends. *Med Image Anal.* **2018**, *47*, 45–67. [[CrossRef](#)]
28. Wang, X.; Ahmad, I.; Javeed, D.; Zaidi, S.A.; Alotaibi, F.M.; Ghoneim, M.E.; Daradkeh, Y.I.; Asghar, J.; Eldin, E.T. Intelligent Hybrid Deep Learning Model for Breast Cancer Detection. *Electronics* **2022**, *11*, 2767. [[CrossRef](#)]
29. Hirra, I.; Ahmad, M.; Hussain, A.; Ashraf, M.U.; Saeed, I.A.; Qadri, S.F.; Alghamdi, A.M.; Alfakeeh, A.S. Breast cancer classification from histopathological images using patch-based deep learning modeling. *IEEE Access* **2021**, *9*, 24273–24287. [[CrossRef](#)]
30. Li, Y.; Wu, J.; Wu, Q. Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. *IEEE Access* **2019**, *7*, 21400–21408. [[CrossRef](#)]
31. Neuner, C.; Coras, R.; Blümcke, I.; Popp, A.; Schläffer, S.M.; Wirries, A.; Buchfelder, M.; Jabari, S. A Whole-Slide Image Managing Library Based on Fastai for Deep Learning in the Context of Histopathology: Two Use-Cases Explained. *Appl. Sci.* **2021**, *12*, 13. [[CrossRef](#)]
32. Lo, C.M.; Wu, Y.H.; Li, Y.C.J.; Lee, C.C. Computer-Aided Bacillus Detection in Whole-Slide Pathological Images Using a Deep Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 4059. [[CrossRef](#)]
33. Pedersen, A.; Smistad, E.; Rise, T.V.; Dale, V.G.; Pettersen, H.S.; Nordmo, T.A.S.; Bouget, D.; Reinertsen, I.; Valla, M. H2G-Net: A multi-resolution refinement approach for segmentation of breast cancer region in gigapixel histopathological images. *Front. Med.* **2022**, *9*, 971873. [[CrossRef](#)]
34. Fisher, R.A. *Contribution to Mathematical Statistics*; John Wiley and Sons: New York, NY, USA, 1950.
35. Rutherford, A. *Introducing ANOVA and ANCOVA: A GLM Approach*; Introducing Statistical Methods Series; John Wiley & Sons: Hoboken, NJ, USA, 2001.
36. Turner, J.; Thayer, J. *Introduction to Analysis of Variance: Design, Analysis & Interpretation*; Sage: New York, NY, USA, 2001.
37. Montgomery, D.C. *Design and Analysis of Experiments*; Wiley: Hoboken, NJ, USA, 1984.
38. Ahmedt-Aristizabal, D.; Armin, M.A.; Denman, S.; Fookes, C.; Petersson, L. A survey on graph-based deep learning for computational histopathology. *Comput. Med. Imaging Graph.* **2021**, *2021*, 102027. [[CrossRef](#)] [[PubMed](#)]
39. Hong, R.; Liu, W.; DeLair, D.; Razavian, N.; Fenyö, D. Predicting endometrial cancer subtypes and molecular features from histopathology images using multi-resolution deep learning models. *Cell Rep. Med.* **2021**, *2*, 100400. [[CrossRef](#)] [[PubMed](#)]
40. Wang, S.; Yang, D.M.; Rong, R.; Zhan, X.; Xiao, G. Pathology image analysis using segmentation deep learning algorithms. *Am. J. Pathol.* **2019**, *189*, 1686–1698. [[CrossRef](#)]

41. Angayarkanni, S.P. Hybrid Convolution Neural Network in Classification of Cancer in Histopathology Images. *J. Digit. Imaging* **2022**, *35*, 248–257. [[CrossRef](#)]
42. Bnoui, N.; Amor, H.B.; Rezik, I.; Rhim, M.S.; Solaiman, B.; Amara, N.E.B. Boosting CNN Learning by Ensemble Image Preprocessing Methods for Cervical Cancer Segmentation. In Proceedings of the 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), Monastir, Tunisia, 22–25 March 2021; pp. 264–269. [[CrossRef](#)]
43. Swathi, C.; Anoop, B.; Dhas, D.A.S.; Sanker, S.P. Comparison of different image preprocessing methods used for retinal fundus images. 2017 Conference on Emerging Devices and Smart Systems (ICEDSS), Tiruchengode, India, 3–4 March 2017; pp. 175–179.
44. Hong, R.; Fenyö, D. Deep Learning and Its Applications in Computational Pathology. *BioMedInformatics* **2022**, *2*, 159–168. [[CrossRef](#)]
45. Davri, A.; Birbas, E.; Kanavos, T.; Ntritsos, G.; Giannakeas, N.; Tzallas, A.T.; Batistatou, A. Deep Learning on Histopathological Images for Colorectal Cancer Diagnosis: A Systematic Review. *Diagnostics* **2022**, *12*, 837. [[CrossRef](#)]
46. Dimitriou, N.; Arandjelović, O.; Caie, P.D. Deep Learning for Whole Slide Image Analysis: An Overview. *Front. Med.* **2019**, *6*, 264. [[CrossRef](#)]
47. Saednia, K.; Tran, W.T.; Sadeghi-Naini, A. A Cascaded Deep Learning Framework for Segmentation of Nuclei in Digital Histology Images. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Glasgow, UK, 11–15 July 2022.
48. Umer, M.J.; Sharif, M.; Kadry, S.; Alharbi, A. Multi-Class Classification of Breast Cancer Using 6B-Net with Deep Feature Fusion and Selection Method. *J. Pers. Med.* **2022**, *12*, 683. [[CrossRef](#)]
49. Breast Cancer Dataset. 2016. Available online: <https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images> (accessed on 29 September 2022).
50. Öztürk, Ş.; Akdemir, B. Effects of Histopathological Image Pre-processing on Convolutional Neural Networks. *Procedia Comput. Sci.* **2018**, *132*, 396–403. [[CrossRef](#)]
51. Salvi, M.; Acharya, U.R.; Molinari, F.; Meiburger, K.M. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput. Biol. Med.* **2021**, *128*, 104129. [[CrossRef](#)] [[PubMed](#)]
52. Danelljan, M.; Robinson, A.; Khan, F.S.; Felsberg, M. Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking. In *Computer Vision – ECCV 2016*; Springer International Publishing: Berlin, Germany, 2016; pp. 472–488.
53. Chaki, J.; Dey, N. *A Beginner's Guide to Image Preprocessing Techniques*; CRC Press: Boca Raton, FL, USA, 2018.
54. Pillow Library. 2022. Available online: <https://pillow.readthedocs.io/en/stable/> (accessed on 29 September 2022).
55. Bhuiyan, M.R.; Abdullah, J. Detection on Cell Cancer Using the Deep Transfer Learning and Histogram Based Image Focus Quality Assessment. *Sensors* **2022**, *22*, 7007. [[CrossRef](#)] [[PubMed](#)]
56. Celebi, T.; Shaye, I.; El-Saleh, A.A.; Ali, S.; Roslee, M. Histogram Equalization for Grayscale Images and Comparison with OpenCV Library. In Proceedings of the 2021 IEEE 15th Malaysia International Conference on Communication (MICC), Malaysia, 1–2 December 2021; pp. 92–97. [[CrossRef](#)]
57. Yoon, H.; Han, Y.; Hahn, H. Image contrast enhancement based sub-histogram equalization technique without over-equalization noise. *Int. J. Electr. Comput. Eng.* **2009**, *3*, 189–195.
58. Sheet, D.; Garud, H.; Suveer, A.; Mahadevappa, M.; Chatterjee, J. Brightness preserving dynamic fuzzy histogram equalization. *IEEE Trans. Consum. Electron.* **2010**, *56*, 2475–2480. [[CrossRef](#)]
59. Magudeeswaran, V.; Ravichandran, C. Fuzzy logic-based histogram equalization for image contrast enhancement. *Math. Probl. Eng.* **2013**, *2013*, 891864. [[CrossRef](#)]
60. Vuong, N. 2020 Available online: <https://www.kaggle.com/code/nguyenvlm/fuzzy-logic-image-contrast-enhancement> (accessed on 29 September 2022).
61. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
62. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.0486.
63. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Seattle, WA, USA, 27–30 June 2016; pp.770–778. [[CrossRef](#)]
64. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269. [[CrossRef](#)]
65. Talo, M. Automated classification of histopathology images using transfer learning. *Artif. Intell. Med.* **2019**, *101*, 101743. [[CrossRef](#)]
66. Buddhavarapu, V.G.; J, A.A.J. An experimental study on classification of thyroid histopathology images using transfer learning. *Pattern Recognit. Lett.* **2020**, *140*, 1–9. [[CrossRef](#)]
67. Hameed, Z.; Zahia, S.; Garcia-Zapirain, B.; Aguirre, J.J.; Vanegas, A.M. Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models. *Sensors* **2020**, *20*, 4373. [[CrossRef](#)] [[PubMed](#)]
68. Srinidhi, C.L.; Ciga, O.; Martel, A.L. Deep neural network models for computational histopathology: A survey. *Med Image Anal.* **2021**, *67*, 101813. [[CrossRef](#)] [[PubMed](#)]

69. Kassani, S.H.; Kassani, P.H.; Wesolowski, M.J.; Schneider, K.A.; Deters, R. Deep transfer learning based model for colorectal cancer histopathology segmentation: A comparative study of deep pre-trained models. *Int. J. Med. Inf.* **2022**, *159*, 104669. [[CrossRef](#)]
70. Hameed, Z.; Garcia-Zapirain, B.; Aguirre, J.J.; Isaza-Ruget, M.A. Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network. *Sci. Rep.* **2022**, *12*, 15600. [[CrossRef](#)] [[PubMed](#)]
71. Abbasniya, M.R.; Sheikholeslamzadeh, S.A.; Nasiri, H.; Emami, S. Classification of Breast Tumors Based on Histopathology Images Using Deep Features and Ensemble of Gradient Boosting Methods. *Comput. Electr. Eng.* **2022**, *103*, 108382. [[CrossRef](#)]
72. Srinivasu, P.N.; SivaSai, J.G.; Ijaz, M.F.; Bhoi, A.K.; Kim, W.; Kang, J.J. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. *Sensors* **2021**, *21*, 2852. [[CrossRef](#)] [[PubMed](#)]
73. Wang, W.; Li, Y.; Zou, T.; Wang, X.; You, J.; Luo, Y. A Novel Image Classification Approach via Dense-MobileNet Models. *Mob. Inf. Syst.* **2020**, *2020*, 7602384. [[CrossRef](#)]
74. Wen, L.; Li, X.; Gao, L. A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Comput. Appl.* **2019**, *32*, 6111–6124. [[CrossRef](#)]
75. Panda, M.K.; Sharma, A.; Bajpai, V.; Subudhi, B.N.; Thangaraj, V.; Jakhetiya, V. Encoder and decoder network with ResNet-50 and global average feature pooling for local change detection. *Comput. Vis. Image Underst.* **2022**, *222*, 103501. [[CrossRef](#)]
76. Nandhini, S.; Ashokkumar, K. An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Comput. Appl.* **2022**, *34*, 5513–5534. [[CrossRef](#)]
77. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [[CrossRef](#)] [[PubMed](#)]
78. Rawat, W.; Wang, Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Comput.* **2017**, *29*, 2352–2449. [[CrossRef](#)]
79. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A Comprehensive Survey on Transfer Learning. *Proc. IEEE* **2021**, *109*, 43–76. [[CrossRef](#)]
80. Narkhede, S. Understanding auc-roc curve. *Towards Data Sci.* **2018**, *26*, 220–227.