ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

INTRODUCTORY ALGORITHM FOR EDGE DETECTION IN MAMMOGRAM IMAGES: A DIDACTIC APPROACH

ALGORITMO INTRODUTÓRIO PARA DETECÇÃO DE BORDAS EM IMAGENS DE MAMOGRAFIA: UMA ABORDAGEM DIDÁTICA

BERNARDO TERNUS DE ABREU
https://orcid.org/0000-0003-4081-9590 / http://lattes.cnpq.br/9234153102615950 / bernardoternus@gmail.com
Universidade do Vale do Rio dos Sinos, São Leopoldo, Rio Grande do Sul

CAIO FRANCISCO TERNUS DE ABREU
https://orcid.org/0000-0003-2496-6638 / http://lattes.cnpq.br/0550917796395985/ caio.ternus@gmail.com
Universidade do Vale do Rio dos Sinos, São Leopoldo, Rio Grande do Sul

Recebido em: 09/03/2025
Aprovado em: 09/07/2025
Publicado em: 19/07/2025

RESUMO

Este artigo tem como objetivo apresentar um algoritmo didático para detecção de contornos em imagens de mamografias. O algoritmo utiliza métodos de processamento de imagens e recebe uma imagem de entrada de um exame de mamografia, retornando uma imagem de saída filtrada, com a identificação dos nódulos existentes. Foram utilizadas imagens de um repositório acadêmico para fins educacionais. O algoritmo foi capaz de identificar contornos na imagem de saída e apresentou alta precisão (0.9536), indicando poucos falsos positivos. A acurácia foi moderada (0.6023), mas o desempenho foi limitado por valores baixos de recall (0.2270), F1-Score (0.0718) e IoU (0.0374). Os dados indicam que, embora acerte quando detecta, o algoritmo não detectou uma parcela dos contornos reais. Apesar disso, o algoritmo pode ser aprimorado como uma ferramenta didática, com vistas ao aprendizado de conceitos que estão presentes em determinadas tecnologias.
Palavras-chave: contorno, detecção, mamografias

ABSTRACT

This paper aims to present a didactic algorithm for contour detection in mammogram images. The algorithm uses image processing methods and receives an input image from a mammogram exam, returning a filtered output image, with the identification of existing nodules. Images from an academic repository were used for educational purposes. The algorithm was able to identify contours in the output image and presented high precision (0.9536), indicating few false positives. Accuracy was moderate (0.6023), but performance was limited by low recall (0.2270), F1-Score (0.0718) and IoU (0.0374) values. The data indicate that, although it is correct when detecting, the algorithm did not detect a portion of the real contours. Despite this, the algorithm can be improved as a didactic tool, with a view to learning concepts that are present in certain technologies.
Keywords: contour, detection, mammograms

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

INTRODUCTORY ALGORITHM FOR EDGE DETECTION IN MAMMOGRAM IMAGES: A DIDACTIC APPROACH

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

1 INTRODUCTION

Breast cancer is the second most common type of cancer in the world, after lung cancer, and the most common type in women. Early detection and treatment are factors that increase the expectation of a cure in cases of cancer, and the most accessible and used techniques are mammography and self-examination. According to Kaplan, Malmgren, Atwood and Calip (2015), between the 1990s and 2010, there was a 34% reduction in mortality among women in the United States due to breast cancer. The data indicate that the reduction is directly associated with the implementation of mammography (Kaplan, Malmgren, Atwood, Calip, 2015, p. 2553).

In the 1980s, mammograms became established as a screening test. Statistics from 1979 indicated that in a study of 20,000 women, 31 percent of the cancers found by physical examination were tiny lesions. Current studies indicate that it is easier to detect tiny lesions using mammograms than through physical examination or self-examination by the patient.

A first relevant question about cancer detection by medical imaging would be: What imaging techniques are used for tumor detection? In response, there are different breast screening methods or techniques that have been developed in recent decades. They are: Breast Ultrasound (BU), Mammography, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Thermography, and Biopsy (Rezaei, 2021, p. 2). Although there are other technologies for imaging diagnosis, they are not used for breast cancer due to the complexity of the disease.

A second research question is: What are the essential factors for detecting breast cancer by imaging? In this case, it is necessary to consider that breast tissues may have clusters of different intensities, being composed of fibroglandular, adipose and pectoral muscle tissue. The abnormal area in a lesion involves calcification or poorly defined masses (Rezaei, 2021, p. 3). Generally, a malignant tumor refers to cancerous cells and a benign tumor indicates non-cancerous tissue.

Malignant cancer is represented by clustered distributions of microcalcifications with linear branching pattern involving more than 3 microcalcifications together, with 6 diameters less than 0.5 mm, while benign cancer is considered as individual microcalcifications. Microcalcifications are the main factor to be taken into account in the detection of breast cancer in mammograms. As a result, the analysis and interpretation of microcalcifications are fundamental for the detection and reduction of mortality rates (Rezaei, 2021, p. 4).

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

One of the challenges in the correct detection of malignant tumors is false negative diagnoses, usually linked to noise in images and low quality of the data obtained on microcalcifications in a given patient from whom images were collected. Because of this, technologies that support data collection and interpretation can be useful for healthcare teams.

Currently, new technologies are being tested to support medical decision-making. Image processing algorithms are being improved to identify lesions in exams.The Food and Drug Administration (FDA) recently approved the use of deep learning technologies to support medical decision-making in terms of automatic segmentation of mammogram images. This approval represents an important milestone, as it scientifically validates the use of systems based on artificial intelligence (AI) in image-assisted diagnosis, especially in digital breast tomosynthesis (DBT), also known as 3D mammography.

Among the approved systems is ProFound AI Version 3.0, from the company iCAD, authorized in 2021. This model demonstrated significant clinical improvements, such as increased specificity and sensitivity, in addition to reducing the time to read exams by up to 40%. Another notable system is Saige-Q, developed by DeepHealth, also approved in 2021, which allows automatic screening of mammograms with support for both 2D and 3D exams. It was followed by Saige-DX, launched in 2022, with a focus on reducing false positives and unnecessary recalls.

In addition, MammoScreen, from the French company Therapixel, received FDA authorization for use in breast tomosynthesis, presenting clinical benefits such as a higher cancer detection rate and a lower incidence of false alarms. More recently, in 2023, the FDA approved Lunit Insight DBT, further reinforcing the presence of automatic segmentation and classification algorithms in supporting radiological diagnosis.

These technologies stand out for providing precise segmentation of suspicious regions in images, in addition to generating heat maps, detection boxes and risk scores. With this, radiologists can make faster and more informed decisions, increasing the efficiency and accuracy of diagnosis.

This article will discuss introductory aspects of mammography image processing, focusing on segmentation. Modern methods such as machine learning and deep learning will not be applied, but rather simpler resources used to demonstrate mammographic image processing, with a view to introducing concepts.

This article presents introductory concepts, from a didactic perspective, on image processing for medical images of cancer. However, the technologies used in the industry are much more complex than the algorithm presented, which aims only presenting initial concepts in the field.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

INTRODUCTORY ALGORITHM FOR EDGE DETECTION IN MAMMOGRAM IMAGES: A DIDACTIC APPROACH

Bernardo Ternus de Abreu
Caio Francisco Ternus de Abreu

From an algorithm that filters images and identifies contours, it is possible to develop a basis for techniques more commonly found in studies in the field of computing, which in the last decade have established convolutional neural networks as a very recurrent technique.

In order for the student to learn what a convolutional neural network is in the context of images, however, it may be advisable to start with simpler algorithms, and this project can be understood as an example of an initial programming project involving medical images for undergraduate biomedical engineering course. Students who have already graduated or have more experience in the field could resort to neural networks and other methods.

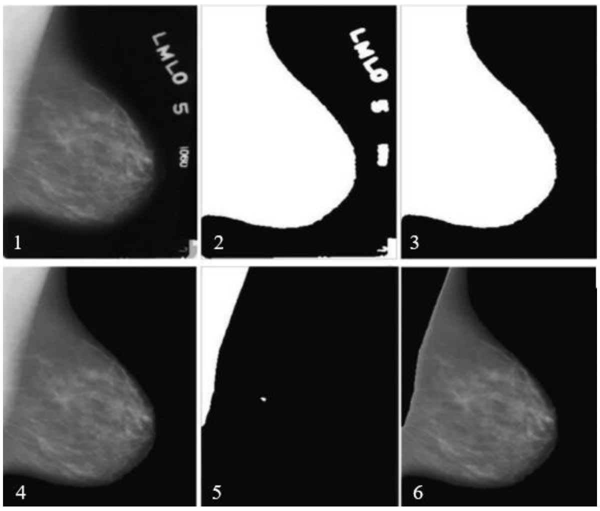
2 THEORETICAL BACKGROUND

2.1 BREAST CANCER AND MEDICAL IMAGING

Cancer originates from changes in the DNA code, the chromosome code, from mutations in genes. It occurs in multiple types of tissues and organs, with a variety of pathological consequences for different organs and tissues, so that it represents a complex disease at different levels: genetic, histological, pathological, in terms of prognosis and therapy.

To remove muscle tissue from a mammogram in an automated way using a computer, an algorithm that performs a set of steps can be considered. First, the original image 1 is binarized (image 2). Then, a “mask” is applied (image 3). The mask is a matrix used to modify or analyze an image, and is applied to the original image through a convolution operation, in which each pixel of the output image is calculated based on the neighboring pixels of the original image, weighted by the values of the mask. After removing the label or caption (image 4), the pectoral muscle is identified (image 5) and the pectoral muscle is removed (image 6). Image 6 is the result of the preprocessing, so that the image can then move on to the segmentation stage.

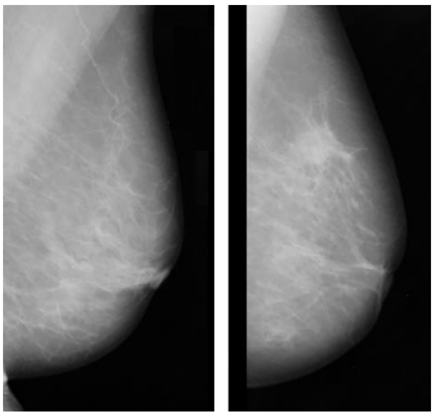
Figure 1 - Steps for Removing Muscle Tissue from a Mammogram



Source: adapted from Rezaei et al (2021, p. 5)

Mammography images allow the visualization of nodules inside the breast of patients. Figure 2 shows two images, one with a benign tumor on the left, and another with a malignant tumor or cancer on the right. In the mammogram on the right, it is possible to see a mass with an amorphous shape in the upper central portion of the breast.

Figure 2 – Examples of benign and malignant tumors



Source: adapted from Jaamour et al (2023, p. 2)

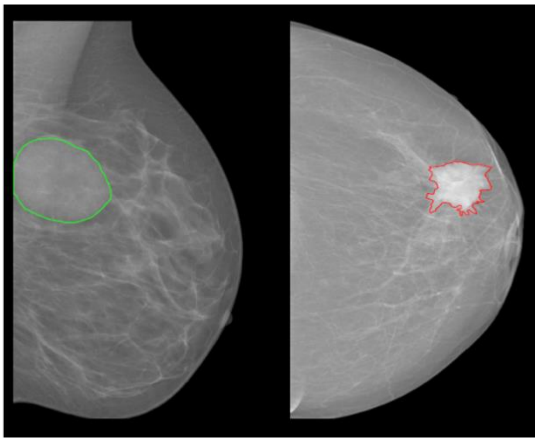
In figure 3, it is possible to visualize this difference with a contour enhancement algorithm, in which the image on the left has a benign tumor, composed of a mass with a regular and uniform shape, while the lesion on the right has a misshapen shape, characterizing typical cancer cell growth.

Figure 3 – Tumor lesions with uniform and misshapen contours

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

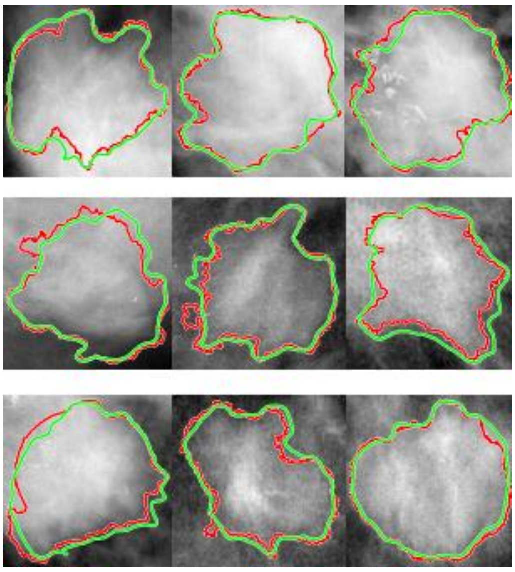
BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU



Source: adapted from Li, Chen, Nailon, Davies, Laurenson (2022, p. 1)

To illustrate Computer Aided Diagnosis (CAD) techniques, an illustration of the output of an article by Li, Chen, Nailon, Davies, Laurenson (2022) can be used. In the image, the algorithm returns images with the green outline indicating the contours. The red outline was performed by the pathologist.

Figure 4 – Lesion delineation performed by: Human versus Computer



Source: adapted from Li, Chen, Nailon, Davies, Laurenson (2022, p. 6)

In the literature, when the true outline of the lesion is demarcated by a team of pathologists and serves as a reference for training an algorithm, the stipulated contour is called “Ground Truth”. From this outline, the model can be evaluated in a comparative manner, using the output trace of the identified contour.

2.2 MAMMOGRAPHY

Mammography is an established technique for early detection of breast cancer. This article focuses on this technique because it is widely used due to its accessibility and low cost. In general, mammography exams are performed annually for patients over 40 years of age, and are performed using low-dose X-rays. Compared to other methods such as magnetic resonance imaging, it has lower accuracy but is more cost-effective. Because it is a low-cost method, it can be included in state policies to assist a large number of people. In terms of a comparison of techniques for breast cancer diagnosis, it can be mentioned 7 different ways:

1) Self-examination: is a procedure that can be performed monthly in every age. It is an inexpensive and non-invasive procedure (Ansar and Raza, 2023, p. 5).

2) Clinical examination of the breasts: is a procedure that can be performed every 3 years and it is used for diagnosis. If there is any family history, it is better to go to the clinic for a regular check-up.

3) Mammography: is indicated for women with 40 years or older, for annual exam. It Is used for scanning and diagnosis. Advantages of the method includes: Low-dose X-rays. Sensitivity is 90% for breasts with fatty tissue. Early detection is possible even before the examination is felt by patients. In addition, the cost is low. As disadvantage, it can be mentioned that Sensitivity is low for dense breasts. The cancer-free rate on mammograms is 25% due to low intensity and cancer detection is difficult if the patient that has dense breasts (Ansar and Raza, 2023, p. 5).

4) Ultrasound: is an exam that is indicated depending on the medical condition. It is Indicated when the breast is dense. It is a broad and effective complement to mammography, and can differentiate between solid and cystic lesions. As disadvantage, The accuracy of the image and its interpretation depend on the person involved in the examination. It is not possible to detect cancers less than 1 mm in diameter.

5) Magnetic Resonance Imaging (MRI): is performed under medical prescription. MRI is 99% sensitive for detecting breast cancer. It is recommended for patients with a family history of breast cancer. However, it has 10 times the cost compared to mammography.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

6) Biopsy: is performed if cancer is diagnosed. Image-guided breast biopsy performed with MRI or ultrasound. As an advantage, while core needle biopsy provides faster clinical results than surgical excision. As an disadvantage, the treatment process is invasive and the patient may suffer from breast pain, stress, swelling, fatigue and peeling skin (Ansar and Raza, 2023, p. 5).

Mammography is well-established as an imaging modality used for breast cancer screening. It is a non-invasive technique that uses low-dose X-rays to generate high-resolution images of breast tissue. The technique operates on the principle of differential X-ray attenuation: breast tissue is compressed between two plates and a beam of low-dose X-rays is directed through the breast to create an image.

Digital mammography has replaced film-screen mammography, resulting in improved image quality and reduced radiation dose. Digital breast tomosynthesis (DBT), a 3D mammography technique, has increased breast cancer detection rates and reduced false positives. Automated breast ultrasound (ABUS) is another imaging modality used. Mammography is effective in screening for breast cancer and may reduce breast cancer mortality rates, especially in women aged 50 to 74. Additional screening with MRI or ultrasound may be recommended for women at higher risk of breast cancer, such as those with a family history or genetic predisposition (Wang, 2024, p. 2).

Several leading companies and research groups have made significant advances in the past decade. For example, Hologic's Genius 3D mammography technology provides high-resolution 3D images, increasing detection rates and reducing false positives. However, this comes at the expense of increased radiation exposure and higher costs compared to traditional mammography. Other developments include contrast-enhanced spectral mammography (CESM) from GE Healthcare and Siemens Healthineers, which combines mammography with contrast-enhanced imaging to improve diagnostic accuracy (Wang, 2024, p. 2).

Artificial intelligence tools developed by companies such as iCAD and ScreenPoint Medical have been used to improve the interpretation of mammograms, leading to earlier detection of breast cancer. Gamma Medica and Dilon Technologies have introduced new breast imaging technologies, such as molecular breast imaging and breast-specific gamma imaging, which use different types of radiation to provide more detailed images of breast tissue.

The University of Chicago has made advances in contrast-enhanced mammography (CEM), which is more accurate at detecting invasive breast cancers than traditional mammography alone. CEM provides detailed images of breast tissue without ionizing radiation, although it is not widely available and may not be covered by insurance plans.

Several DL architectures, including convolutional neural networks (CNN), transfer learning (TL), ensemble learning (EL), and attention-based methods, have been developed for various applications in mammography. These applications include breast cancer image detection, classification, segmentation, restoration, and enhancement, and computer-aided diagnosis (CAD) systems.

Artificial intelligence-based models, such as machine learning algorithms, can analyze medical imaging datasets and patient characteristics to identify breast cancer or predict the risk of developing breast cancer. Machine learning algorithms can extract quantitative features from medical images, such as mammograms or ultrasound images, through radiomics. AI based prediction models can incorporate multiple cancer risk factors, including genetics, lifestyle, and environmental factors, to establish personalized imaging and treatment plans. In recent years, deep learning algorithms have emerged as promising AI tools to increase the accuracy and efficiency of breast cancer detection (Wang, 2024, p. 2).

2.3 PRE-PROCESSING AND FILTERS

Image acquisition is the first step in medical image processing algorithms, where the patient's image is obtained. After acquisition, pre-processing begins, which involves applying filters and other techniques to improve image quality by removing noise or distortion. This may include the use of smoothing filters, such as the Gaussian filter, or enhancement techniques, such as histogram equalization. Next stage is segmentation, a crucial step where the image is divided into regions of interest, such as organs or lesions, to facilitate analysis.

Image preprocessing is performed to improve image quality, reduce noise and remove unwanted information from images. Image filters are operations that alter or improve the appearance of an image by applying mathematical transformations to the pixels. They are used to smooth, enhance, or detect specific image features, such as edges or textures. In general, low-pass and high-pass filters are used for noise reduction.

Low-pass and high-pass filters are used in pre-processing to manipulate the frequencies present in the image: the low-pass filter, such as Gaussian blur, smooths the image by attenuating high-frequency components, such as noise and small abrupt variations between pixels, preserving homogeneous regions. The high-pass filter, on the other hand, enhances edges and fine details, amplifying rapid variations in intensity, although it can also enhance noise.

Histogram equalization, adaptive histogram equalization and Contrast-Limited Adaptive Histogram Equalization (CLAHE) are used for image enhancement. Histogram equalization redistributes brightness levels globally, expanding the dynamic range of the image and improving overall contrast, but it also tends to increase noise.

Adaptive histogram equalization (AHE) applies the same concept locally, dividing the image into regions to adjust contrast independently, which can result in localized improvement but amplifies noise in homogeneous areas. Contrast-Limited Adaptive Histogram Equalization (CLAHE), on the other hand, improves on adaptive histogram equalization (AHE). It limits contrast gain per region to avoid strong highlights in uniform regions of the image.

Image resizing and data augmentation are also performed. Data augmentation is necessary for deep learning-based systems because pre-trained CNN architectures have different input layers. Due to this, images are rotated or cropped (Thakur, Kumar, Kumar, 2024, p. 35933).

Filters can be spatial, such as those that act directly on the image, or they can be frequency filters, when the image is first transformed to the frequency domain, using the Fourier transform. Generally, this process is done by the discrete Fourier transform, so that the signals are filtered in the frequency domain and then the image is transformed back to the spatial domain (Santos, 2011, p. 22).

The Median Filter is one of the filters found in the literature review. It is a non-linear filter used in image processing to remove noise. Each pixel in an image is changed to have the median value of its neighbors. Another filter found in the literature is the Wiener filter, which consists of a linear filter used in image processing to improve image quality and eliminate noise.

The Wiener filter was designed to reduce the mean square error between the filtered and original images, and was formulated based on the idea that the signal and the noise are uncorrelated and operate in the frequency domain. Since Y(f) is the Fourier transform of the noisy signal and H(f) is the frequency response of the transfer function, Sn(f) and Sx(f) are, respectively, the power spectral density (PSD) of the noise and the original signal.

In the context of data augmentation in computer vision, flips and random crops are fundamental techniques to increase the diversity of the training set, improving the generalization of models. The Mirroring or Flips technique involves inverting an image horizontally or vertically. This technique is useful when the orientation of the object is not fixed, such as in object recognition or facial expression recognition tasks. By applying random flips during training, the model learns to recognize features regardless of orientation, reducing the risk of overfitting and improving robustness.

The Random Crops technique consists of randomly cropping a region of the original image, resizing it to a fixed size. This technique is effective for simulating variations in the position and scale of objects, forcing the model to learn certain representations, being useful in object detection and classification tasks.

2.4 SEGMENTATION TECHNIQUES

Image segmentation consists of separating or delimiting regions corresponding to objects or structures represented in an image. It is an important step in extracting features from the elements of an image and can be performed using different methods, the most common being segmentation by discontinuity and similarity. In discontinuity, abrupt changes or edges in an image are sought, while in similarity, pixels are grouped according to their similarity of features.

Segmentation algorithms extract portions of images based on characteristics such as: color, texture or intensity. They define sub-regions in the images to be able to perform the action. The most commonly used segmentation methods are: 1) Threshold-based segmentation: which calculates one or more limits according to a rule or function, assigning the pixels to different classes according to these limits; 2) Edge-based segmentation: identifies sudden changes in intensity, being useful in detecting the edges or corners of an image; 3) Region-based segmentation: partitions are made in the image based on similarity criteria; 4) Cluster-based segmentation: The image is divided into groups based on similarity of pixel color, intensity, or location. An example of an algorithm of this type is K-means (Rezaei, 2021, p. 5).

Thresholding converts images into binaries by separating background and object intensities, and is simple and fast, but sensitive to noise and variable illumination. Morphological techniques use operations such as erosion and dilation to refine binarized shapes, to remove noise and fill gaps. The zero-crossing method identifies contours in which the intensity gradient changes sign, and is indicated for smooth edges.

The watershed technique models the image as a topographic relief and segments it by "flooding" up to ridge marks, but it can lead to over-segmentation. Clustering techniques, such as K-Means or Fuzzy C-Means, group pixels by similarity of intensity or texture, being effective for homogeneous regions. Classifier-based segmentation methods employ trained models, such as SVM and RF, to decide pixel by pixel or region by region.

In turn, the Markov random field models spatial relationships and prioritizes regularity in segmentation, adapting to probabilistic models. Active contours adjust deformable curves to fit gradient-guided lesion edges. Level sets adjust implicit contours according to minimized energy functions. Texture-based segmentation methods use descriptors, such as local binary patterns (LBP) or Gabor transform, to distinguish tumor regions from normal regions based on their textural features.

The Otsu method is one of the most widely used techniques, because from a gray image, it determines the ideal value of a threshold that separates the elements of the background and foreground of the image into two clusters, assigning the color white or black to each of them. In cases of bimodal histogram images, the technique allows to adequately divide the regions into a single value. The Otsu concept consists of finding the threshold 1 between two pixel values that divide an image, minimizing the variance in each group and increasing the variance between groups. The first group is formed by values from 0 to 1 and the second group from 1 to G, where G is the largest possible pixel intensity. The method would test all values of 1 in the interval [1, G, 1] in order to maximize the variance equation between groups.

In the region growing method, the technique starts with a known seed point and expands to neighboring pixels that share similar characteristics, such as intensity or texture, thus forming a region that may correspond to a tumor. Edge segmentation focuses on detecting abrupt discontinuities in pixel intensities, indicative of edges or contours of objects, such as the boundaries of a tumor. Active contour-based segmentation, known as "snakes", uses curves that move across the image to find object boundaries, adapting to the characteristics of the tumor. There is also histogram-based segmentation, which uses the distribution of pixel intensities in the image to identify different types of tissue, such as cancerous areas versus normal areas.

The multilevel thresholding technique based on the Kapur method, combined with the Shell Game Optimization (SGO) algorithm, known as OKMT-SGO, is an efficient approach for image segmentation, especially in preprocessing tasks. Multilevel thresholding seeks to divide an image into several distinct regions based on different pixel intensity ranges, allowing for a more refined analysis of visual structures.

The Kapur method is used to select these thresholds based on maximizing the image entropy, that is, identifying the separation points that offer the greatest amount of information possible between the segmented regions.To find the optimal set of thresholds that maximizes this

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

entropy, the SGO algorithm is used, a technique that simulates a solution search process by exploring the search space. By applying OKMT-SGO before feeding convolutional neural networks, it is possible to significantly improve the quality of the inputs (Kavitha et al, 2022, p. 115).

There are several segmentation techniques, as it can be seen. In the case of hybrid approaches, they combine multiple of these techniques. There are also methods based on machine learning and deep learning, such as U-Net or CNN networks, which learn visual patterns and automatically segment lesions with high precision, especially with large annotated databases. These methods are recommended because they require less manual input throughout the process, once the model is established.

In recent years, libraries have been developed in Python programming languages and their respective documentation on how to use them for machine learning and deep learning applications. Tools such as TensorFlow, PyTorch, Keras, among others, offer a robust infrastructure for the creation, training and validation of artificial intelligence models, allowing the development of algorithms with some pre-made functions.

The TensorFlow library enables the construction of neural network architectures, allowing the development of multilayer models for deep learning. In the context of medical or histopathological imaging applications, integration with the OpenSlide library enables efficient handling of whole-slide images, enabling cropping into regions of interest and conversion of images to model-compatible formats. The combination of these tools enables automated workflows for image analysis, with the potential to impact computer-aided diagnosis.

For Ali and Hamed, edge detection-based segmentation is a method generally employed with gray histogram and gradient-based method. Zadeh, Kazenouni, and Haddadnia used parabolic transformation for region of interest (ROI) segmentation. For this purpose, contours were detected using a logarithmic method. The initial results were contaminated with noise, which was removed with a Gaussian filter before edge detection (Zadeh, Kazenouni, Haddadnia, 2011).

2.5 MACHINE LEARNING AND DEEP LEARNING

Over the years, classical segmentation methods have been partly replaced or accompanied by modern methods, which use computational resources from the field of artificial intelligence.

Artificial Neural Networks are formed by an architecture that has: input layer, output layer and intermediate or hidden layers between them. The input layer delivers the patterns to the Artificial

Neural Network, while the intermediate layers perform the processing of much of the project. The intermediate stage is responsible for feature extraction, while the output layer concludes and presents the final results. The topology defines the way in which neurons are connected, and they can be fully interconnected or partially connected.

Goodfellow, Bengio, and Courville (2016) define deep learning as a subfield of artificial intelligence that is concerned with the study of programs that learn knowledge from experience using models graph-based computational algorithms, where concepts are learned in layered subgraphs. Deep learning composes representations of more abstract concepts from less abstract representations, where the representation is a mathematical object that succinctly describes another set of objects, such as images.

In 1943, McCulloch and Pitts published a machine model based on the concept of a human brain. Later, in 1949, with the work of Hebb, a basic model of a self-organizing network was presented. In 1958, Rosenblatt proposed the Perceptron model, a fundamental work for artificial neural networks. In the following decades, neural networks gained the ability to be trained from data and received a wide range of architectures and applications. In a 1989 article, LeCun and collaborators demonstrated backpropagation training and presented architectural changes that were important for the next convolutional neural networks, such as AlexNet (Lima, 2023, p. 21).

Some aspects of neurons that must be taken into account are: (1) connection weights, (2) summing junction, and (3) activation function. At each input of the system there is a (1) distinct or equivalent weight, which can be modified to make an adjustment in data processing. The (2) summing junction is responsible for summing the inputs of the neurons that are weighted by the corresponding connection weights and for generating an induced local field, which is known as activation potential. Finally, the (3) activation function is responsible for limiting the activation potential to ensure that the amplitude has a finite value, depending on the activation function used (Peixoto, 2017, p. 10).

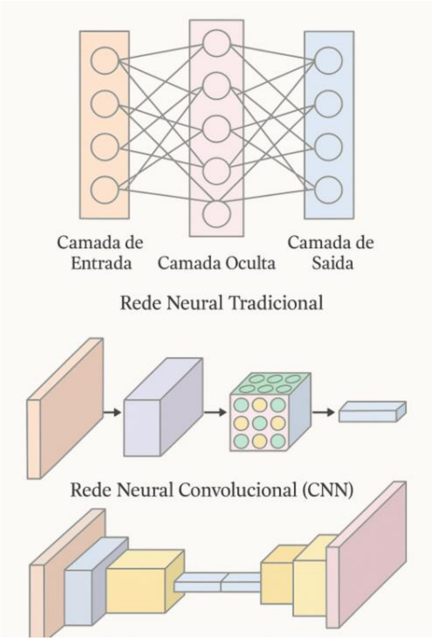
At the low level of processing or transformation, the input and output of the process are images. In this case, applications such as contrast enhancement and color correction can be cited. At a medium level of transformation, there are processes that transform the image in a more comprehensive way, being able to separate regions, in the segmentation process, or extract information from the images. In this case, the output is a new image. At the highest level of processing, there is an understanding of the context of the image and the algorithm can perform

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

cognitive functions of the past information, which resembles the process of the biological mind (Kelsch, 2022, p. 17).

Figure 5 – Characteristics of neural networks



Source: adapted from Haque and Neubert (2020)

From sets of images that make up a database, the network is trained by adjusting the weights of the connections between the layers to minimize the difference between the model's prediction and the correct label, using a loss function. During training, the network learns to identify complex patterns in images, becoming capable of classifying new images based on the acquired knowledge.

Convolutional Networks are multilayer neural networks that are composed of two modules: (1) automatic feature extractor and (2) classifier. The feature extraction module consists of four operations: 1) convolution, 2) activation, 3) sampling, and 4) normalization. If the network architecture has a convolution in one or more of its layers, the network can be called convolutional (Peixoto, 2017, p. 13).

Convolutional Neural Networks learn a hierarchical representation of features, using strategies such as local fields, shared weights (convolution) and sampling (pooling). The neuron of a layer connected to a set of neurons of a previous layer forms the local receptive field of this neuron. It is through this local receptive field that each neuron of the first hidden layer can detect features.

In Neural Networks, each neuron is connected to all neurons of the previous layer. With this, both pixels that are close and those that are in distant regions of the image are treated in the same way. An artificial neural network with locally connected layers does not have this problem, since each neuron is connected to a subset of neurons of the previous layer, which are in the neighborhood (Peixoto, 2017, p. 16).

In Convolution, the intensity of a given pixel is calculated based on the intensity of its neighbors. Each pixel in the output image can be considered as a weighted sum of nearby pixels in the input image, where the weights and the size of the region to be averaged are defined by the convolution kernel.

The activation operation is responsible for mapping the resulting map of the convolution operation in order to threshold the output neuron. Activation functions can be: Rectified Linear Unit (ReLU), Rectified Linear (ReL), Logistic Sigmoid and Hyperbolic Tangent. The “rectified linear units” (ReLUs) function learns filters using a supervised algorithm.

The pooling operation function is used to reduce the size of the spatial representation, with a view to reducing the number of parameters and network computation. It reduces the sensitivity of the output map to shape distortion changes, and selects invariant features that can improve generalization performance. The normalization operation creates competition between the maps and attempts to support generalization in order to increase the quality of the convolutional neural network in terms of the output feature vector (Peixoto, 2017, p. 21).

In terms of machine learning architecture, it can be divided into three types of learning: (1) supervised, (2) unsupervised, and (3) reinforcement. In supervised learning (1), there are: (1.1) classification models; (1.2) regression models. In unsupervised learning (2), there are: (2.1) hierarchical models, (2.2) clustering models. Finally, in reinforcement learning (3), there are: (3.1) recommendation models and (3.2) reward models (Melo, 2021, p. 4).

In supervised learning, the computer knows the data and the decisions that need to be made, and it saves the information in memory locations for application to unknown data. In contrast, in unsupervised learning, the computer has no knowledge of the target classification, and it is necessary for the computer to learn about the behavior of the data over time. Finally, in reinforcement learning, the computer learns the classification criterion through trial and error, from information that is provided during the execution of the algorithm (Melo, 2021, p. 1).

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

To build a machine learning algorithm, two sets of data are required: training and testing. From the training data set, it is possible to adjust parameters and detect patterns. In this first stage, information is provided that can be applied in the next stage, testing. The test data set is not intended to be used for parameter adjustment, but is used to test models made with the training data, in order to validate whether the model that was developed from the training data is effective in making predictions.

In the parameter adjustment stage, adjustments are made so that the algorithm can detect patterns in the data. There are two situations, either underfitting or overfitting, that indicate that an algorithm has not been trained correctly: underfitting and overfitting, both expressions related to the training “fit”. When the model is not able to adjust parameters because it cannot detect patterns in the data based on the information provided to the algorithm during training, in this case there is an underfitting scenario. This occurs when the model’s performance is poor during training, so that the algorithm is not able to identify relationships between the variables. As a result, the algorithm loses its practical usefulness.

With the inclusion of neural networks as a segmentation tool, it is important to understand how they work for segmentation. A relevant question could be: How can neural network architecture perform primary segmentation?

Preprocessed mammography images and their corresponding processed masks can be fed into neural network architecture, such as a U-Net CNN. The architecture captures the complex features of mammography images and balances the fine details.

Skip connections help preserve spatial information at different scales, which is relevant for breast image segmentation. The first step usually consists of a binary segmentation, indicating whether there is an abnormality or whether the tissue is normal. The machine learning or deep learning model seeks to capture the local and global context in the image. The output is a primary segmentation image with regions of interest highlighted, which serves as input for subsequent refinement.

The input mammographic image undergoes a preprocessing stage, in which features are extracted and the database is subsequently divided. Part of the data is used for a training database, while part of the data is used for a test database. The proposed model receives data from the training dataset, as well as existing models. From the proposed model, be it a U-Net, SegNet or another architecture, an initial segmentation is performed. With the existing models, segmentation is also performed with a view to comparing the results. A more refined segmentation is performed from the

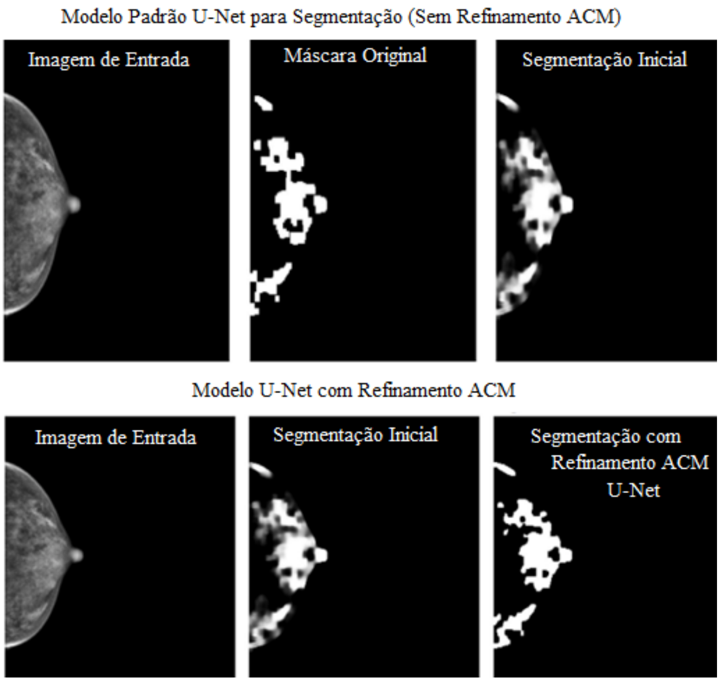
ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

proposed model, with a view to comparing the output results with the results of the existing models, considering the metrics conventionally used for evaluation (Nour, Boufama; 2025, p. 9).

Figure 6 illustrates the results of applying image segmentation to an input image. It shows the output image of a network with and without a specific refinement step. The illustration is adapted from the article by Nour and Boufama (2025, p. 15-16).

Figure 6 – Results of a neural network for segmentation of mammographic images



Source: adapted from Nour and Boufama (2025, p. 15-16).

Region of interest (ROI) segmentation is a step in the automated analysis process of mammography images. It can be considered a classification task because it classifies each pixel in the images of the dataset, either from the region of interest or from the background of the image. The input of the segmentation step consists of mammography images and the output is the images of the ROI or regions of interest.

The ROI images are “masked” with the original mammography images to prepare them as input for the classification phase. An example of segmentation can be using the modified U-Net architecture. The modified U-Net segmentation is composed of the encoder and decoder network. A traditional CNN contains semantic information and less spatial information, which is called the encoder part. However, spatial information is also important for the segmentation of semantic information.

The specific information from the decoder part is fed into the U-Net, where the semantic information is extracted from the lowest layer of the U-Net network. The decoder part contains the high-resolution features, these features are extracted from the encoder part. A ReLU with a data normalization and standardization step is employed in the paper by Salama and Aly (2021, p. 4703).

U-Net is a convolutional neural network architecture that has an image compression path and an image expansion path, a configuration that became known as "encoding-decoding". In the encoder network, a sequence of convolution and pooling layers reduces the size of the input image and extracts features at different scales. U-Net networks have a “U”-shaped structure that ensures that each pixel in the output contains all the information from the input. The input and output data in the training and test sets must be the same size, or undergo some processing to adjust the dimensions.

The concept of transformer was studied in the article "Attention Is All You Need" (2017), which introduces an architecture based on transformers. The paper introduces a new deep learning model architecture that breaks away from conventional approaches based on recurrent neural networks (RNNs) and convolutional neural networks (CNNs) for sequence modeling tasks. Instead of relying on sequential mechanisms to process input token by token, Transformer relies solely on attention mechanisms, which allows for greater parallelism and better ability to capture long-range dependencies between sequence elements.

The core component of the architecture is the scaled dot-product attention mechanism, which computes a distribution of attention weights between pairs of positions in the sequence by taking the dot product of the query and key vectors, scaled by the square root of the dimensionality of the key vectors, followed by the application of the softmax function. These weights are then used to compute a linear combination of the value vectors associated with each key. This operation allows the model to extract contextually sensitive representations for each position in the sequence, taking into account its relationship with all other positions.

To increase the model’s ability to capture multiple semantic relations in parallel, the Transformer implements the concept of multi-head attention, where input vectors are projected into distinct subspaces and processed by multiple independent attention heads. Each head implements its own scaled attention mechanism, allowing the model to learn different types of patterns and interactions in different representation subspaces. The outputs from these heads are then concatenated and re-projected to produce the final output of the attention layer. This design increases the flexibility and expressiveness of the model without significantly increasing the computational cost, since each head operates in a reduced dimension.

The complete Transformer architecture is divided into two symmetric parts: the encoder and the decoder. The encoder is composed of a stack of identical layers, each containing two main sublayers: a multi-head attention sublayer applied to the input and a fully connected feedforward network, both followed by layer normalization and residual connections operations.

The decoder has a similar structure, but includes an additional masked multi-head attention sublayer, which prevents the model from attending to future positions of the output sequence during autoregressive training, ensuring that the prediction of a token depends only on previous tokens. In this case, the decoder includes a cross-attention sublayer, which allows it to query the latent states produced by the encoder, establishing a bridge between input and output.

Since the architecture is invariant to the order of tokens, since the attention mechanisms are permutable, the Transformer introduces positional embeddings (positional encodings) to incorporate information about the relative or absolute position of tokens in the sequence. These embeddings are vectors added to the input embeddings and can be learned or defined by fixed trigonometric functions. This addition allows the model to preserve the sequential structure of natural language during the attention process.

The complete elimination of recurrence and convolution, combined with extensive use of attention, makes Transformer highly parallel and efficient for training on large data sets. Its ability to capture long-range dependencies directly, as well as its modular and scalable architecture, have made it the basis for several subsequent state-of-the-art models in natural language processing tasks, such as BERT and GPT. Transformer represented a paradigm shift from sequential learning to an attention-only approach, with significant implications for computational efficiency and the quality of learned linguistic representations.

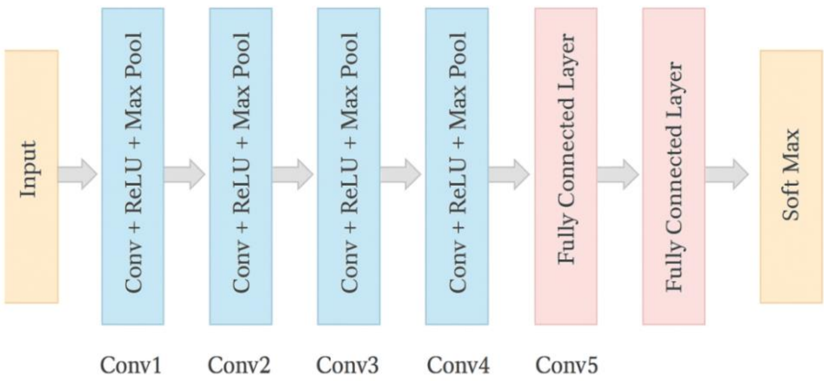
The paper "Learning representations by back-propagating errors" (1986), by Rumelhart, Hinton and Williams, formalized the backpropagation algorithm, essential for training deep neural networks by adjusting their weights based on the error in the output. Although the mathematical concept of backpropagation had existed before, with roots dating back to the 1970s, this work was responsible for popularizing it in the machine learning community, by demonstrating in a clear, didactic and convincing way how multilayer neural networks could be trained efficiently.

In the paper, the authors presented a practical approach to adjusting the weights of artificial neural networks by calculating the gradient of the error with respect to each weight using the chain rule. This allowed the output error to be “backward propagated” through the layers of the network, updating the parameters in an efficient and supervised manner. Prior to this work, many attempts to train neural networks with more than one layer failed due to the lack of systematic and stable methods for adjusting the internal weights, the so-called “depth problem.”

The main contribution of the paper was to show that networks with multiple hidden layers could learn internal representations (or “distributed representations”) useful for complex tasks such as pattern recognition, information encoding, and decision making. In other words, it was not necessary to explicitly program rules to solve cognitive tasks — networks could learn to extract their own useful representations from data, provided they were trained with enough examples and with an effective error correction mechanism, such as backpropagation.

Decades later, the paper "ImageNet Classification with Deep Convolutional Neural Networks" (2012) by Krizhevsky, Sutskever, and Hinton demonstrated the power of convolutional neural networks (CNNs) in computer vision tasks, marking the beginning of the modern era of deep learning with the use of GPUs to train large-scale models. This method became known as AlexNet and introduced the use of techniques such as the ReLU activation function and training on GPUs.

Figure 7 – AlexNet Architecture



Source: adapted from Mohapatra, Muduly, Mohanty, Ravindra, Mohanty (2022, p. 299).

The AlexNet architecture consists of eight weighted layers, the first five of which are convolutional layers and the last three are fully connected layers. The output of the last layer is fed into the Softmax function.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

The series of convolution, pooling and fully connected layer in Alex Net classifier extracts the features from mammograms, reduces the dimension of the extracted feature and finally classifies images into benign, malignant and normal classes (Mohapatra, Muduly, Mohanty, Ravindra, Mohanty, 2022, p. 299).

In 2014, two papers were particularly important. The first, “Sequence to Sequence Learning with Neural Networks” introduced the seq2seq model with an encoder-decoder based on RNNs, a basis for machine translation and sequence tasks. The second, “Generative Adversarial Nets”, introduced GANs, networks that learn to generate new and realistic data through a game between two networks: generator and discriminator, opening the door to modern generative AI.

In 2014, it was introduced the Residual Networks or ResNet, which enabled the training of extremely deep networks by facilitating gradient flow. ResNet presented a structure with residual connections between layers, allowing the training of models with good results for images. In 2015, some studies began to highlight the role of attention as a way to improve seq2seq models, allowing them to focus on relevant parts of the input.

The question that drives engineers and scientists who use these methods is: how can we build computer programs that improve over time based on input data? Deep learning is an approach that has been used in this direction. The concept of deep learning initially emerged from research into artificial neural networks. The most basic example of model is known as a "feedforward deep network" or multilayer perceptron, also known as a Multilayer Perceptron (MLP).

Over time, it was realized that the perceptron created by Rosenblatt had a limitation: the model was only capable of solving linearly separable models. Linearly separable models are those that can be separated by a straight line in a 2D coordinate system or by a plane in a multidimensional system. According to Minsky and Selfridge (1961), the model was not capable of solving a simple non-linearly separable problem, such as the XOR logic gate. Because of this, it was not possible to use the perceptron to make general abstractions (Haykin, 1994). Later, Minsky and Papert (1969) proved through mathematical functions that Rosenblatt's perceptron is incapable of making any global generalization based on the examples learned.

After the publication of Minsky and Papert's book, a series of doubts arose for years about the computational power of the perceptron and neural networks, which persisted until the mid-1980s. In 1986, Rumelhart published a method using a supervised algorithm known as backpropagation. With this work, he showed the importance of internal representations in neural networks, also called hidden layers.

Backpropagation consists of a method for training artificial neural networks, used in conjunction with optimization methods, such as gradient descent. With this technique, it is possible to train networks that are capable of generalizing any type of function, making it possible to solve non-linear problems. After the publication of this work, neural networks with hidden layers began to be tested in different applications.

A neural network can be trained with input and output data from a data set, and the error can be calculated for each iteration. The error is the difference between the expected result and the result generated by the neural network using weights randomly generated at the beginning. Knowing the error, it is possible to adjust the weights of the neural network to empirically reduce the error. The point at which the algorithm converges occurs when adjusting the weights in the network no longer affects the overall error, thus stabilizing the network error.

The purpose of the pooling layer is to reduce the resolution of the input activation map and concentrate the information. As the features are detected by the algorithm, the exact position of the feature becomes less important for the final decision. Pooling layers can be of two types: "max pooling" and "average pooling". In the max pooling layer, the maximum value is selected, while in the "average pooling" layer an average of all inputs is returned.

The greater the number of multiplications performed throughout the backpropagation process, the smaller the gradients become. In these cases, the multiplication elements have a modulus less than one. According to Costa (2023, p. 23), in some situations, the gradients reach zero, preventing the initial layers from being updated. When this occurs, the phenomenon is called “vanishing gradients”. Skip connections seek to reduce this problem by ensuring the continuous flow of gradients from the initial layers to the final layers.

The encoder-decoder architecture, also known as encoder-decoder, is used in deep learning tasks that involve sequence transformation. The architecture is characterized by two main components: the encoder, responsible for encoding the input into a vector representation, and the decoder, which uses this representation to generate the output. Both components are implemented as recurrent neural networks (RNNs), but modern versions employ other architectures. These are: LSTM (Long Short-Term Memory), Gated Recurrent Unit (GRU) or even Transformers.

The encoder processes the input sequence element by element, updating its internal state at each time step. At the end of the sequence, the final state of the encoder is assumed to be a compact representation, or context vector, that contains the input information. This vector is then passed to the decoder, which uses it as the initial condition for generating the output sequence. The

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

decoder operates sequentially, producing one element of the output at a time. The prediction at a time step depends on the previous predictions.

Although effective, this architecture has limitations, especially when dealing with long sequences. Reliance on a single context vector can lead to information loss, which can negatively affect the quality of the output. To mitigate this problem, techniques such as "attention mechanisms" have been introduced. These mechanisms allow the decoder to directly access all intermediate states of the encoder, dynamically weighting the most relevant parts of the input at each step of generating the output. This advancement culminated in the development of the Transformer, an architecture that completely abandons RNNs and relies exclusively on attention, increasing parallelization and the ability to model long-term dependencies.

The softmax layer is used in the output layer of the network to classify the input into different categories. The softmax layer transforms a vector of input values into a probability distribution, where each element in the output represents the probability of the input belonging to one of the possible classes. This transformation is performed by applying a normalized exponential function to the input values. The softmax layer configuration is useful in classification tasks, as it allows the convolutional neural network to assign probabilities to each possible class. This fact is important for assessing the model's confidence in the result of its predictions

Some works use the U-Net architecture for the segmentation of medical images. For image preprocessing, some authors use Fully Convolution Network (FCN) networks within the structure of a CNN. The use of U-net networks augmented by dilation layers and squeeze and extraction units can be found in the literature (Lima, 2023, p. 27).

In some cases, a preprocessing step is used to remove edges. CNN convolutional networks of the encoder-decoder type with "squeeze" and "excitation" blocks are used. To reduce false positives, the literature also presents studies that use image post-processing techniques, such as using the extraction and selection of image patches in the training phase of an FCN, to reduce complexity and increase sample variability (Lima, 2023, p. 29).

Some deep learning models use a ResUnet++ architecture and apply the Conditional Random Field (CRF) method to extract useful geometric features from the model, such as shape and region connectivity. Subsequently, a data augmentation technique is used in the test base, called Test-Time Augmentation (TTA). In the literature, cases of including attention blocks in the direction of the ResUnet++ encoder can be found. Some authors modify the UNET network by adding attention layers and using the Tversky Loss function in the training phase.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

A methodology used in deep learning involves the DT-WpCNN network, which has a pooling mechanism based on multi-resolution analysis, built from a wavelet transform. With this method, it is possible to reduce the dimensionality of the feature maps of the previous convolutional layer while preserving the information carried. In this network, a technique called LGWe-LSM is applied to the input images, which allows the generation of active contour maps of the tumor region. By combining the two results: DT-WpCNN and LGWe-LSM, it is possible to reduce the number of false positives. In the literature, the combination of a U-Net with a MobileNetV2 is used to extract image features. The U-Net architecture can be trained with different structures.

First, with an EfficientNet-B4 and then with an EfficientNet-B5. A loss function can be presented to solve the problem of data imbalance. In the literature, it is possible to find works that use a CNN SE-Resnext-50 as an encoder and two decoders that work in parallel, where one segments the possible area containing the tumor and the other segments the tumor contour. In addition, a simple U-Net containing two layers can be used for the regression of the result of the first U-Net and also for the result of the second U-Net (Lima, 2023, p. 30).

A framework for tumor segmentation can be used using a pyramid-view Transformer as an encoder for feature extraction. In this case, modules that assist in the segmentation task are used: i) CFM: a module for collecting semantic and location information about tumors through progressive integration. ii) CIM: a module for detecting camouflaged objects, which uses an attention mechanism, reducing incorrect information in the lower features. iii) SAM: this module consists of a graph convolutional layer to mine local data and global semantic features.

As for the image segmentation task, the works show a tendency to use automatic and traditional methods for object segmentation. These are: encoder-decoder methods, such as U-Net and their combinations. A significant portion of the works use preprocessing techniques for input images, in order to correct problems resulting from image acquisition. In addition, post-processing techniques are often used to reduce false positives in the final segmentation result (Lima, 2023, p. 29).

In the literature, the use of the ResNet-101 architecture as the backbone of a Mask R-CNN network can be found. ResNet-101 is a deep neural network architecture from the Residual Networks (ResNets) family, proposed by researchers at Microsoft Research. Its network is composed of 101 trainable layers, and is designed with residual blocks, which allow learning very deep networks without the problem of degradation or loss of performance.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

The innovation of ResNets lies in the introduction of shortcut connections (skip connections), which allow the input signal of a block to be added directly to its output. This approach facilitates gradient propagation during training, making learning more efficient and stable in networks with many layers.

When integrated into architectures such as the encoder-decoder CNN, the SE technique helps to improve the detection of regions of interest. Encoder-decoder architectures have the function of extracting and then reconstructing or segmenting images. The encoder-decoder CNN architecture with residual blocks with SE (squeeze and excitation) has been applied to detect regions containing tumors. Some works use Transformer mechanisms, such as a YOLOv5 architecture with the aid of a Self-Attention mechanism for tumor detection in images. In the feature extraction stage, an Attention mechanism can be added to increase the contribution of channels with more features and reduce the interference of channels with fewer features.

In the preprocessing stage, the training images can be joined into mosaics with the aim of increasing the variability of the images for the model (Lima, 2023, p. 31). Deep learning is based on multilayer computational models, such as CNNs, which learn data representations at various levels of abstraction (Lima, 2023, p. 43).

The success of deep learning-based architectures is due to the model's ability to generalize, even if its training uses different types of heterogeneous images. Instead of being programmed to solve a specific type of problem, deep learning models are able to extract the main characteristics of the object being studied and learn from them, ensuring better performance in the results.

Some works after 2020 indicate a trend, among current mammographic image segmentation methods, towards the use of hybrid architectures (CNN + transformer).

Iqbal and Sharif (2023) proposed a hybrid BTS-ST model, composed of a BTS-ST vision transformer-based network for segmentation and classification of breast tumor images. The primary structure of BTS-ST is based on a dual encoder using U-Net and Swin-Transformer, Spatial Interaction Block (SIB), Feature Compression Block (FCB) and Relationship Aggregation Block (RAB). The BTS-ST model obtained a segmentation accuracy of 0.996, higher than other models such as U-Net (0.984), TransUNet (0.990) and TransFuse (0.992). The hybrid model integrates CNN and transformer and obtained better results than methods based purely on CNNs such as U-Net. The article currently has 82 citations since 2023, indicating its use by authors in the field.

Mohammadi and Livani (2025) proposed a hybrid CNN + transformer model, called Hybrid Transformer U-Net (HTU-net), developed for the segmentation of breast masses in mammograms.

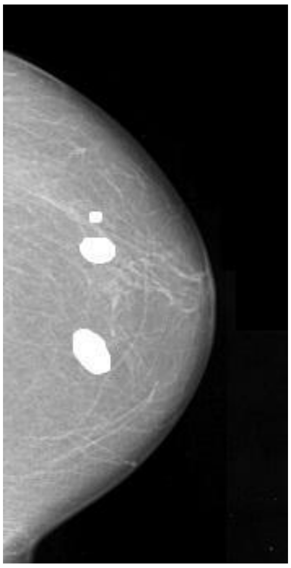
ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

The authors used the following databases: CBIS-DDSM and INBreast. The spatial and channel self-attention mechanisms were improved in the model and integrated into the convolutional layers in HTU-Net, creating a hybrid architecture. The architecture combines the strengths of CNNs and ViTs. The authors introduced a multiscale attention mechanism that further enhances the model's ability to fuse information from different resolutions, improving the decoder's ability to reconstruct fine details in the segmented output. Mohammadi and Livani (2025) improved the model to capture local texture-based features and global contextual information, aiming at a better segmentation result.

3 MATERIALS AND METHODS

The images database comes from a public digital repository at the University of South Florida (http://www.eng.usf.edu/cvprg/). The USF repository is open access and has a database of mammogram images with patients whose identification is not mentioned, which provides security in handling data for educational studies. The code of the developed can be found at the link: https://github.com/beternus/contour-detector/blob/main/contourdetector.py

Figure 9 – First Adapted Mammography Image from the Repository of USF



Source: The authors (2025)

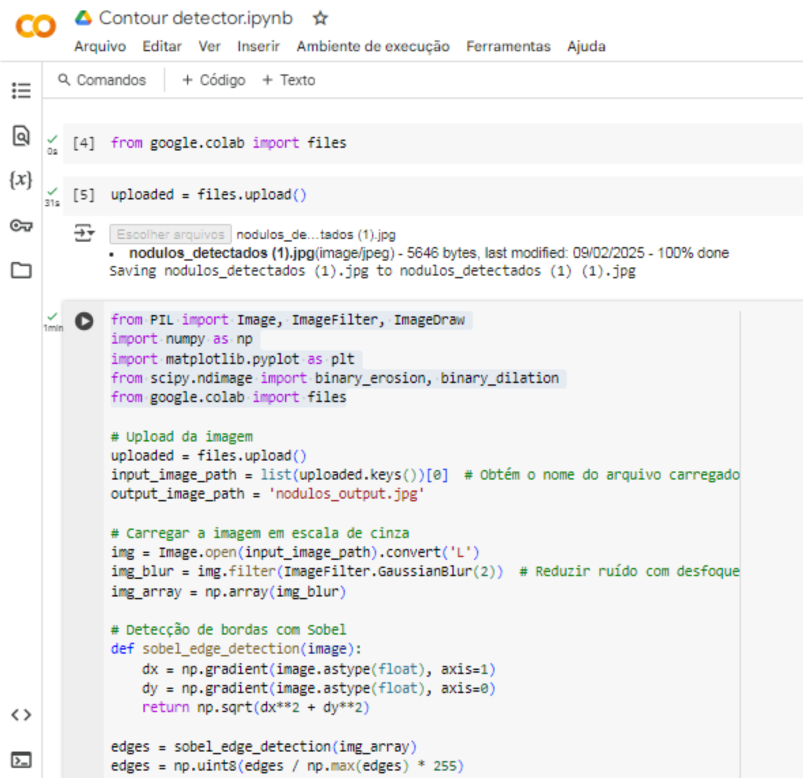
ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

Bernardo Ternus de Abreu
Caio Francisco Ternus de Abreu

The programming language adopted in this project was Python, as it is a language widely used in image processing and computer vision, and because it has well-established libraries for carrying out certain operations. The active development ecosystem of this language has been mainly responsible for the increase in quality for its data analysis and visualization tools for large datasets over the past couple of decades. This language provides access to powerful libraries for storing, manipulating and gaining insights from data, such as Numpy for efficient storage and manipulation of dense data arrays, and Matplotlib for a flexible range of data visualization capabilities. Next, we will mention about the development of the algorithm model for image processing in Python.

First, the authors downloaded some .jpg images available in the USF public repository. After that, two images were modified by manually introducing blank nodules in the images for an initial testing. The compiler used to process the code was Google Colab, an open-source tool that can be run over the internet, without the need to download a specific execution environment. In Google Colab, the image output results are displayed in the interface itself, below the code and during execution, which makes it easier to observe.

Figure 10 – Code execution



ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

Source: the authors (2025)

The implemented code imports libraries for image processing and visualization. The “PIL” (Python Imaging Library, now maintained as “Pillow”) is used for opening, manipulating, and saving image files, with modules like “Image” for image creation and manipulation, “ImageFilter” for applying filters, and “ImageDraw” for drawing shapes or text on images. The library “NumPy” is imported to handle numerical operations, particularly for working with image data as arrays or matrices. Matplotlib's “pyplot” module is used for generating visualizations, including plots and graphs, and it helps in displaying images or results.

The “SciPy” library's “ndimage” submodule provides functions like “binary_erosion” and “binary_dilation”, which are used for morphological operations on binary images, shrinking or expanding bright regions. It is used the “google.colab.files” module for file upload and download functionality within a Google Colab environment, allowing user to interact with files stored on local machine or in the cloud.

The code first uses “files.upload()” from the google.colab library to prompt the user to upload a file in the Google Colab environment. The uploaded file is stored in a position of memory, where the keys are the file names. The code then retrieves the name of the uploaded file by accessing the first key in the dictionary using “list(uploaded.keys())[0]”, which is stored in the variable “input_image_path”. This variable now contains the path to the uploaded file. The “output_image_path” variable is defined as “nodulos_output.jpg”, which will be used later to store the path for saving the processed output image.

After the image has been loaded, its name defined and the name of the output image defined, the image is converted to grayscale. The code opens the image file specified by “input_image_path” using the “Image.open()” method from the PIL library. It then converts the image to grayscale using the “.convert('L')” method, which simplifies the image to shades of gray, reducing the complexity for further processing. Next, the image undergoes a Gaussian blur filter using “img.filter(ImageFilter.GaussianBlur(2))” to reduce noise, where the argument 2 specifies the radius of the blur effect. The blurred image is stored in the “img_blur” variable. Finally, the blurred image is converted into a NumPy array using “np.array(img_blur)”, which allows for numerical manipulation and processing of the image data in array form.

After the image has been converted to grayscale and converted to numeric form, the edge detection step occurs. The code defines a function “sobel_edge_detection” that performs edge

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

detection using the Sobel operator. It calculates the gradients of the image in both the x-direction (dx) and the y-direction (dy) by applying “np.gradient()” to the image array. The gradients are computed by converting the image to a float type using “astype(float)”.

The function then combines the gradients using the formula “np.sqrt(dx**2 + dy**2)” to calculate the magnitude of the edges at each point in the image. After defining the function, it is applied to the “img_array” to detect edges, and the resulting array is normalized to the range of 0 to 255 by dividing by the maximum value and scaling. The edges array is then thresholded, where values less than 100 are set to 0 (indicating no edge) and values greater than or equal to 100 are set to 255 (indicating an edge), creating a binary image where edges are highlighted.

After edge detection, some adjustments are made to the image. The code defines two functions, “dilatacao” and “erosao”, which apply morphological operations to a binary image. The “dilatacao” function uses “binary_dilation” from the “scipy.ndimage” library to expand the bright regions in the binary image, using a 3x3 matrix of ones as the structuring element. The “erosao” function uses “binary_erosion”, also from “scipy.ndimage”, to shrink the bright regions of the binary image, again with a 3x3 matrix of ones as the structuring element. These operations are applied sequentially to the edges array: first, erosion is applied to the edges with the “erosao” function, and then dilation is performed on the eroded edges using the “dilatacao” function, resulting in an image where the edges are processed and modified by both shrinking and expanding operations.

The code defines a function “find_contours” that extracts contours (connected components) from a binary image, where edges are represented by white pixels (255). It uses a breadth-first search (BFS) algorithm to find and trace these contours. The function starts by initializing an empty list, contours, to store the detected contours and a visited array to keep track of already processed pixels. Inside the “bfs” function, a queue is initialized with the starting pixel coordinates, and the algorithm explores all 8-connected neighboring pixels in all directions to identify the full contour of a connected region. If a contour meets a minimum size requirement that is set by “min_area” it is added to the contours list. The outer loop iterates over every pixel in the edges image, and for each unvisited white pixel, it calls “bfs” to find the contour. Finally, the function returns a list of all detected contours, which are groups of connected edge pixels.

The next lines creates an overlay image by making a copy of the original image “img” using the “.copy()” method, so that the original image remains unchanged. It then initializes a drawing object with “ImageDraw.Draw(overlay)” to allow for drawing on the copied image. The loop iterates over each detected contour in the contours list, and for each contour, the “draw.polygon()” method

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

is used to draw a red outline around the contour on the overlay image. This method connects the points of the contour to form a polygon, and the “outline=red" argument ensures that the contour is drawn in red color.

The code begins by setting up a figure for displaying multiple images in a grid using “plt.figure(figsize=(12, 8))”, which defines the size of the figure. Then, the “plt.subplot()” function is used to create a 2x3 grid of subplots, and each subplot is populated with a different image. In the first subplot, “plt.imshow(img, cmap='gray')” displays the original image in grayscale, followed by “plt.title("Imagem Original")” to set the title of the subplot. The second, third, and fourth subplots follow a similar pattern, showing the blurred image, detected edges, and the result after erosion, each with a corresponding title indicating the processing step. After placing all images in their respective subplots, “plt.tight_layout()” ensures that the layout is adjusted so the subplots don’t overlap. The “plt.show()” command finally renders the figure with all the subplots. After displaying the images, the code saves the image with detected contours to the file path “output_image_path” and prints a message confirming the saved image. Finally, the image is made available for download using “files.download(output_image_path)”, allowing the user to download the processed result.

The project flowchart can be seen in Figure 11, which presents the different stages of the algorithm that was described. The code architecture is simple and does not involve more complex resources such as neural networks. Its purpose is didactic and a presentation of introductory concepts in signal processing, such as filters. In this case, the image is modified from a gray scale.

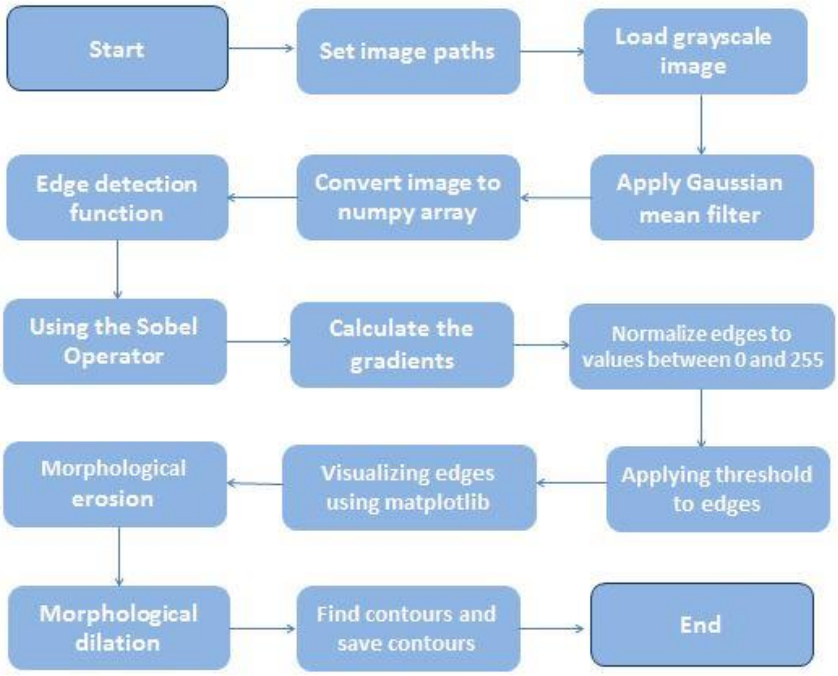
Figure 11 – Flowchart of Algorithm

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

INTRODUCTORY ALGORITHM FOR EDGE DETECTION IN MAMMOGRAM IMAGES: A DIDACTIC APPROACH

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU



Source: The authors (2025)

The algorithm allows students who are having their first contact with the field to introduce some concepts of image progression. Certain practices, such as converting an image into a numerical matrix, can be approached with practice, which uses Python resources such as Numpy. In addition, the student can visualize the application of different types of edge detection techniques, because although the Sobel edge detector was used, other techniques can be used, and it is possible to show how they work in a table. In the Sobel detector, pixel intensity gradients are calculated in the horizontal (X-axis) and vertical (Y-axis) directions to identify where significant brightness variations occur in the image. These variations indicate the edges of the objects present. The output of edge detection is a matrix that represents the magnitude of the gradient, where higher values correspond to stronger edges. This method can be applied to images for other purposes, not only medical images, so that the response of the algorithm can be visualized and whether the results are suitable for the specific purpose.

To improve detection accuracy, the algorithm normalizes the gradient magnitude values to the range 0 to 255 and applies a threshold to eliminate weak edges that are likely caused by noise. In this case, any value less than 100 is discarded (set to 0) and values greater than or equal to 100 are kept as 255. This results in a binary image, where the edges are clearly highlighted. This

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

Bernardo Ternus de Abreu
Caio Francisco Ternus de Abreu

operation with values greater than or equal to 100 justifies the conversion of the image data to a numerical scale, and can be used to discuss engineering strategies that are used with data sets to better manipulate them.

The algorithm allows learning about morphological operations with images, since it performs two fundamental operations: erosion and dilation. Morphological operations are techniques used in image processing, mainly for binary images, where pixels have values of 0 or 1, representing black and white, respectively. They modify or extract shapes and structures from the image based on the shape of the objects present. Erosion is an operation that reduces objects in the image by removing pixels from the edges of the objects, making them smaller. Dilation, on the other hand, expands objects by adding pixels to their edges, and can be useful for filling holes or connecting nearby objects. In the algorithm in question, erosion is applied to reduce fine details and eliminate noise by removing the thinnest edge pixels, while dilation, performed after erosion, serves to restore the main edges, making them thicker and more continuous.

Other fundamental operations are: “opening”, “closing” and “gradient”. “Opening” combines erosion followed by dilation, and is useful for removing small objects or noise from the image, while maintaining the structure of larger objects. “Closing” is the opposite of opening, consisting of dilation followed by erosion, and is used to close holes or gaps within objects, smoothing their edges. The morphological “gradient”, which is the difference between dilation and erosion, helps to highlight the edges of objects, highlighting the changes in intensity between the background and the objects. These operations are performed with the help of a "structuring element", which is a mask or kernel that defines the shape and size of the change in the image, which can be a simple matrix such as 3x3, or other larger and more complex shapes. They are often used in conjunction with other image processing techniques to improve tasks such as segmentation, edge detection, object recognition and shape analysis.

Edge detection is an important step in image processing algorithms. In the algorithm used, a breadth-first search (BFS) is used to identify and extract the connected regions of pixels that correspond to the detected edges. During the search, the edges are collected and the area of each edge is checked. If the area of an edge is greater than a minimum value (50 pixels by default), it is considered relevant and added to the list of detected edges. This step defines how many nodes are identified in the image, and is also based on numerical values of the vector. It is described by a line of code such as: "def find_contours(edges, min_area=50)".

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

The matplotlib library has a function to represent output images with graphic details, with numerical indication on the sides. This resource can be used by engineering students to represent output data in different applications. The output shows images after applying morphological, erosion and dilation operations. However, in a demonstration context, other filters can be included, such as the Gaussian filter, which assigns more weight to the central region and partially blurs the lateral portions of an image.

In the data visualization stage, the algorithm uses the "subplots" feature. A subplot is a smaller individual graph or image within a larger figure or canvas, allowing multiple visualizations to be displayed together in an organized manner. The “plt.subplot()” function is used to create these subplots by specifying the number of rows, columns, and the position of the current graph within the grid. For example, “plt.subplot(2, 3, 2)” divides the figure into 2 rows and 3 columns, with the graph being placed in the second position in that grid. This configuration allows side-by-side comparison of different images or graphs within the same window.

4 RESULTS AND DISCUSSION

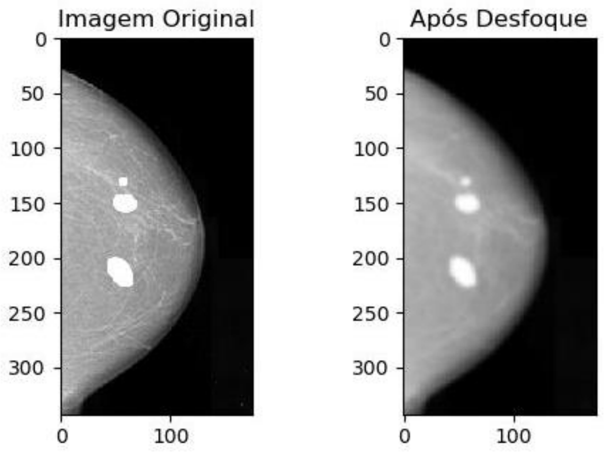
The algorithm was developed for academic/educational purposes in the context of an engineering course and is being improved for educational purposes and not for real/clinical contexts due to its simplicity. The language used was Python, as it is a language widely used in image processing and computer vision, and because it has well-established libraries for carrying out certain operations.

The blurring step was able to smooth the lines of the image to allow for subsequent processing, as can be seen in the image comparison below for the case of the second image, which was the original image that underwent editing to include three nodules.

Figure 12 – Blurred image

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

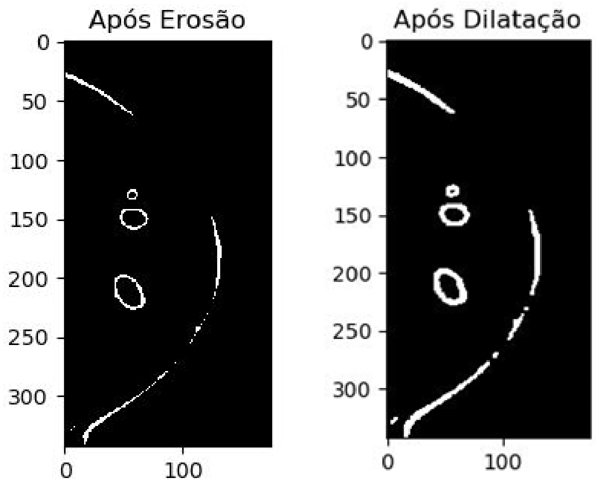
BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU



Source: The authors (2025)

The erosion step allowed the reduction of small signals and eliminated noise, from a line of code that removed thinner edge pixels. The results can be seen in the image below. On the other hand, the dilation was able to restore the main edges, giving characteristics of thickness and continuity to the contours.

Figure 13 – After Erosion vs After Dilation



Source: The authors (2025)

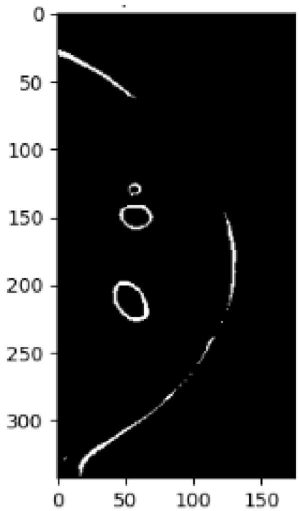
The algorithm's output image highlights the nodules that were present in the patient's breast tissue. Figure 14 shows a black background in the output image, with the tone of the image being quite dark for the most part. In the central portion, the five nodules that had been included in the

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

original image in more or less rounded shapes are identified. The output algorithm highlighted these images. In the middle of the original image there were vessels, resulting from the blood supply to the breast tissue, and which were in white, like the circles drawn. Certain parts of the nodules were not demarcated by the algorithm or the demarcation was attenuated because the structures behind them partially affected the drawing by the algorithm.

Figure 14 – Algorithm Processed Output for First Image



Source: The authors (2025)

To evaluate the algorithm's ability to interpret the data, a second algorithm was developed in Python to compare two images: the initial mammographic image, with the contours of the lesions manually introduced, and the filtered output image, which returns the nodules and the background of the image in black. The algorithm evaluated the results based on the comparison between the output image and the input image. The code for the metrics evaluation algorithm can be found at: https://github.com/beternus/contour-detector/blob/main/metrics.py

Table 1 - Algorithm evaluation through metrics

Metrics	Result			
	Image 1	Image 2	Image 3	Media
Precision	0.9743	0.8926	0.9940	0.9536
Recall (sensitivity)	0.5993	0.0265	0.0551	0.2270

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

Accuracy	0.5993	0.5925	0.6151	0.6023
F1-Score	0.0595	0.0515	0.1044	0.0718
IoU (Jaccard)	0.0307	0.0264	0.0551	0.0374

Source: the authors (2025)

The metrics provide a quantitative way to assess the quality of the segmentation or detection performed by an algorithm, and they are calculated by expressions. Precision is calculated as the number of true positives divided by the total number of pixels predicted as positive by the algorithm, that is, precision = true_positives / (true_positives + false_positives).

Recall (or sensitivity) represents the algorithm's ability to find all the positive pixels from the ground truth and is given by recall = true_positives / (true_positives + false_negatives). Accuracy measures the overall proportion of correct classifications, both positive and negative, and is calculated as accuracy = (true_positives + true_negatives) / (true_positives + false_positives + true_negatives + false_negatives).

The F1-score is the harmonic mean of precision and recall, used to balance these two metrics, and is defined as F1 = 2 * (precision * recall) / (precision + recall). IoU, also known as the Jaccard index, measures the overlap between the prediction and the reference, and is given by IoU = true_positives / (true_positives + false_positives + false_negatives).

The results of the contour identification algorithm metrics present strengths and weaknesses. The main strength is the precision of 0.9536, which indicates that when the algorithm identifies a contour, it is almost always correct. In other words, there are few false positives. However, the other metrics reveal significant limitations. The recall of 0.2270 is extremely low and shows that the algorithm is missing the vast majority of real contours. The F1-Score result is also very low, at 0.0718, which combines precision and recall and, in this case, confirms the imbalance between them. The accuracy, at 0.6023, is moderate, but since this is a problem with imbalanced classes, this metric can be misleading and does not reflect the real performance. Finally, the IoU of 0.0374, measures the overlap between the prediction and the real contour and is also quite low. Low IoU suggests that the detected contours have little intersection with the true contours. In short, the algorithm is able to detect contours, which is indicated by high accuracy, but it is missing almost all the true contours, compromising the detection effectiveness.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

5 FINAL CONSIDERATIONS

The developed algorithm showed the capability to detect contours in mammography images. The algorithm model structure adopted the Sobel method, which is simpler than other techniques such as the Canny algorithm that could be less susceptible to noise. A possibility of future improvement would be the implementation of a dynamic threshold, since the current threshold for edge binarization was fixed at 100 value, which may be ineffective depending on the image in case it has very high or low contrasts. As a possible adjustment, it could be implemented an adaptive threshold, such as the Otsu Method, which could automatically calculate the best threshold value based on the image histogram to improve the detection accuracy.

Another possible approach would be to integrate machine learning to improve the accuracy of edge and contour detection, especially in cases of more complex images or images with significant variations in characteristics. It may be possible to use convolutional neural networks (CNNs), such as U-Net, to segment objects in the image, allowing the detection of regions of interest with much greater accuracy than methods based solely on edges. A neural network could also be trained to automatically classify detected regions of interest, such as nodules in mammograms, increasing the usefulness of the algorithm in medical scenarios. A layered neural network can be used to classify medical images as pathological or healthy through a supervised learning process, where the model is trained with a set of labeled images.

The developed algorithm could be considered an educative image processing resource to familiarize students with image processing topics, considering its low complexity when compared to the cutting-edge algorithms found in the literature review. Considering the increase in articles published since the 2010s involving neural networks for image processing applications, the methodology applied might possess a scientific gap and technical distance from the most recent studies, but can be justified in the didactic and engineering teaching context and purpose.

There is a collective project by several Nations to improve social indicators in the coming decades. The Sustainable Development Goals (SDGs) cover a range of topics, such as eradicating poverty and developing industries and infrastructure. In particular, goal 3 on the list deals with "Health and well-being" and has several targets. One of the targets is to reduce deaths from non-communicable diseases (NCDs) such as diabetes, cancer and cardiovascular diseases. In this sense, research on this topic is justified because it is of collective interest.

The algorithm was able to identify contours in the output image and presented high precision (0.9536), indicating few false positives. Accuracy was moderate (0.6023), but performance was limited by low recall (0.2270), F1-Score (0.0718) and IoU (0.0374) values. The data indicate that, although it is correct when detecting, the algorithm did not detect a portion of the real contours. Despite this, the algorithm can be improved as a didactic tool, with a view to learning concepts that are present in certain technologies.

In a future stage of the project, an algorithm with concepts from a neural network could be developed. From sets of images that make up a database, the network could be trained by adjusting the weights of the connections between the layers to minimize the difference between the model's prediction and the correct label, using a loss function. During training, the network learns to identify complex patterns in images, becoming capable of classifying new images based on the acquired knowledge. For an implementation of this type, it would be necessary a series of studies and tests.

In the field of mammographic image processing, recent articles have shown that the free DDSM and CBIS-DDSM databases are the most widely used for studies. In addition, the CNN U-Net architecture is the most widely used, with more studies in recent years using hybrid CNN + Transformer architectures. Familiarization with more introductory concepts in the field of image processing, in a didactic context, can be a way to enable later learning of more advanced technologies, such as those mentioned above.

6 REFERENCES

Ali, Y.; Hamed, S. Breast Cancer Detection Using Mammogram Images: A Review of Image Processing Techniques. Biosciences Biotechnology Research Asia, 2015.
Ansar, W.; Raza, B. Breast Cancer Segmentation in Mammogram Using Artificial Intelligence and Image Processing: A Systematic Review. Current Chinese Science, 2023.
Costa, N. Aprendizado ativo profundo para classificação de fácies sísmicas. 2023. 91 f. Dissertação (Mestrado em Matemática) - Instituto de Matemática e Estatística - IME, Universidade Federal da Bahia, Salvador (Bahia), 2023.
Daudt, M. Aplicação de aprendizado profundo para detecção de Exoplanetas. Monografia de MBA de Ciência de Dados na Universidade de São Paulo. São Carlos, 2023.
Lima, A. Um método de dois estágios para detecção de pólipos em imagens de colonoscopia usando aprendizado profundo. 2023.156 f. Tese (Programa de Pós-Graduação em Engenharia de Eletricidade/CCET) - Universidade Federal do Maranhão, São Luís, 2023.
Hanahan, D.; Weinberg, R. Hallmarks of cancer: the next generation. Cell, 2011.

ISSN 1983-1838
(DOI): 10.18624/etech.v18i1.1404

INTRODUCTORY ALGORITHM FOR EDGE DETECTION IN MAMMOGRAM IMAGES: A DIDACTIC APPROACH

BERNARDO TERNUS DE ABREU
CAIO FRANCISCO TERNUS DE ABREU

Kelsch, C. Métodos de segmentação de máculas orais em imagens. Monografia de Conclusão de Curso em Engenharia Eletrônica - Faculdade de Engenharia, Universidade do Vale do Rio dos Sinos. São Leopoldo, 2022.
Iqbal, A.; Sharif, M. BTS-ST: Swin transformer network for segmentation and classification of multimodality breast cancer images. Knowledge-Based Systems, 2023.
Jaamour, A.; Myles, C.; Patel, A.; Chen, S.; McMillan, L.; Harris-Birtill, D. A divide and conquer approach to maximise deep learning mammography classification accuracies. PLoS ONE, 2023.
Peixoto, S. Reconhecimento de caracteres em imagens com ruído usando Deep Learning. Dissertação de Metrado apresentada à Universidade Federal de Ouro Preto. Departamento de Computação. Outo Preto, Minas Gerais, 2017.
Li, H.; Chen, D.; Nailon, W.; Davies, M.; Laurenson, D. Dual convolutional neural networks for breast mass segmentation and diagnosis in mammography. IEEE Transactions Medical Imaging, 2022.
Milosevic, M.; Jankovic, D.; Milenkovic, A.; Stojanov, D. Early diagnosis and detection of breast cancer. Technology and Health Care, 2018.
Mohammadi, S.; Livani, M. Enhanced breast mass segmentation in mammograms using a hybrid transformer UNet model. Computers in Biology and Medicine, 2025.
Mohapatra, S. et al. Evaluation of deep learning models for detecting breast cancer using histopathological mammograms Images. Sustainable Operations and Computers, 2022.
Nissar, I.; Alam, S.; Masood, S.; Kashif, M. MOB-CBAM: A dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms. Computer Methods and Programs in Biomedicine, 2024.
Nour, A.; Boufama, B. Hybrid deep learning and active contour approach for enhanced breast lesion segmentation and classification in mammograms. Intelligence-Based Medicine, 2025.
Peixoto, S. Reconhecimento de caracteres em imagens com ruído usando Deep Learning. Dissertação de Metrado apresentada à Universidade Federal de Ouro Preto. Departamento de Computação. Outo Preto, Minas Gerais, 2017.
Rezaei, Z. A review on image-based approaches for breast cancer detection, segmentation, and classific ation. Expert Systems with Applications, 2021.
Salama W, Aly, M. Deep learning in mammography images segmentation and classification: automated CNN approach. Alexandria Engineering Journal.2021.
Wang, L. Mammography with deep learning for breast cancer detection. Frontiers in Oncology, 2024.
Wei, L., Yang, Y., Nishikawa, R. M. Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. Pattern Recognition, 2009.
Zadeh, H.; Kazenouni, I.; Haddadnia, J. Distinguish breast cancer based on thermal in infrared images. Canadian Journal on Image Processing and Computer Vision, 2011.
Zeiser, F.; Costa, C.; Ramos, G.; Bohn, H.; Santos, I.; Roehe, A. DeepBatch: A hybrid deep learning model for interpretable diagnosis of breast cancer in whole-slide images. Expert Systems with Applications, 2021.