Robust Breast cancer Detection using Faster R-CNN Algorithm

Anisa Dian Pratiwi¹, Irma Permata Sari²

Abstract
One of the most common screening tools for breast cancer detection is ultrasound. However, the lack of qualified radiologists causes the diagnosis process to become a challenging task. Deep learning's promising achievement in various computer vision problems inspires us to apply the technology to medical image recognition problems. We propose a detection model based on the Faster R-CNN to detect breast cancer quickly and accurately. We conduct this experiment by collecting breast cancer datasets, conducting pre-processing, training models, and evaluating the model performance. Based on the experiment result, we obtain that this model can detect breast cancer with bounding boxes. In this model, it is possible to detect the bounding box that is more than what it should be, so we applied NMS to eliminate the prediction of the bounding box that is less precise to increase accuracy.

Keywords: Breast Cancer Detection, Deep Learning, Faster R-CNN

1. Introduction

According to the WHO, one of the most frequent female cancers is breast cancer, with 15% yearly deaths [1]. Therefore, timely detection of breast cancer can help people live longer, die more diminutive, and have a better quality of life [2]. In addition, Breast cancer screening is also an effective way of detecting ambiguous breast lesions early. Breast imaging diagnostic, which includes breast MRI, mammography, and breast ultrasound, is a systematic method of breast screening [3].

Due to its pain-free and comfortable operation and excellent real-time performance, ultrasound is among the most commonly used screening technologies for detecting breast cancer. However, because of the ultrasonic instrument's great sensitivity, it is vulnerable to the impact of various tissue of the body and the surroundings, leading to a lot of speckling sound, making it difficult for the doctor to diagnose. In addition, the lack of qualified radiologists can lead to a decrease in diagnostic efficiency, and the missed diagnosis rate is 10-30% [1] [2].

Conventional machine learning (ML) methods require large amounts of manual segmentation annotation data to train and test models for the classification or segmentation of ultrasound images. On the other hand, manual labeling is costly, time-consuming, and labor-intensive, and it significantly raises the cost of system development [4]. Several papers have proposed methods for breast cancer detection. A paper proposed K-Nearest Neighbor (KNN) and Decision Tree to classify breast cancer. After selecting the Principal Component Analysis (PCA) technique, Wisconsin Diagnostic Breast Cancer (WDBC)
dataset verified these two machine learning algorithms. Based on the findings of the experiments, the KNN classifier outperformed the decision tree classifier in the classification of breast cancer [5].

In the problem of improving classification accuracy, a paper suggested Artificial Neural Network (ANN) for breast cancer classification. The Taguchi approach first identifies the number of matching neurons in one of ANN's hidden layers. The model then goes through the training procedure, selecting the correct number of hidden neurons for the hidden layer depending on the Taguchi method's results. Based on the experimental results, this technique can produce the best accuracy in the classification of breast cancer, which is 98.8% [6]. The paper compared the performance of Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) as classifiers in accuracy, sensitivity, and specificity. The model conducts the training process by Mammographic Image Analysis Society (MIAS). Based on the experimental results, the SVM classifier outperformed the KNN classifier with an accuracy of 96%, sensitivity of 92%, and specificity of 100% [7].

Deep learning technology has allowed image recognition to discover target areas in medical images and classify detected target features. Deep learning's detection and classification technique are comparable to the operating procedure used by doctors to determine diagnoses based on ultrasound results. Thus, the approach becomes a new solution to the earlier issues [8]. The current paper proposed Convolutional Neural Network (CNN) and Uniform Experimental Design (UED) to classify breast cancer. UED uses regression analysis to optimize CNN parameters [9][10]. Another study explored a comparative classification of breast MRI tumors using human-engineered radionics, Transfer Learning from Deep Convolutional Neural Network (DCNN), and the Fusion Method [11]. DCNN shows excellent potential for classifying several very various fine-grained objects. Therefore, the further study proposed a deep learning method based on Bilinear Convolutional Neural Network (BCNNs) for the fine category of breast cancer histopathology images [12].

Therefore, we propose Faster R-CNN to detect breast cancer utilize breast images. We establish an effective model to solve the breast cancer detection issue. In breast cancer detection, we present several crucial contributions to this study, particularly in the categorization of breast cancer using learning methods as follows:

1. We introduce a novel technique for detecting breast cancer utilizing the Faster R-CNN algorithm to train the dataset to develop a viable model. We use a dataset of breast ultrasound images to build our model.
2. We build a model to detect breast cancer. This model is a solution to find the location of breast cancer better than traditional techniques.
3. We test a model to achieve high accuracy in detecting breast cancer swiftly and effectively based on the features. We modify several parameters to obtain the best accuracy value to create the best training model.

Organization: The following is a breakdown of the journal's structure: Section II goes deeper into previous discoveries. Section III discusses the study's issue description. Section IV explains the experimental design, including a feature learning algorithm, a dataset, and pre-processing, while section V presents the study's findings as well as an in-depth analysis. Finally, section VI discusses the conclusion.

2. Related Works

Several researchers have recommended studies into developing detection models using classical methods. Based on statistical parameters, a study proposed a cascade algorithm to separate normal, benign, and malignant diseases. Based on the findings of the experiments, the statistical feature technique extracts mammography images to explain the
intensity and distribution attributes using ImageJ [13]. In the problem of breast lump identification, a paper utilized Minimum Spanning Tree (MST) to select the initial cluster center. For predictive probabilities, the intuitive fuzzy c-means clustering identifies abnormalities in breast cancer patients' mammography images and symptoms. Pearson Chi-Square test (x2) at the level significance of 0.05 suggested a strong link between mammography performance and breast cancer clinical symptoms [14].

Another study suggested SVM with several kernel combinations to solve the problem of breast cancer nodule categorization. The model performs the classification process using 22 morphological characteristics derived from the contours of 100 BUS images and reducing the dataset features by a scalar feature selection technique with a correlation. Based on the experiment results, the method can produce accuracy, and the area under the Receiver Operating Characteristic curve (ROC curve) was 96.98% and 0.980, respectively [15]. Another study presented Gaussian Filter and Edge detection techniques to improve image quality. First, wavelet transforms identify first-order features using pre-processed images, and second-order features are retrieved using Gray Level Co-occurrence Matrix (GLCM). Deep Neural Network (DNN), a multilayer supervised classifier, was then used to classify the statistical parameters. According to the experiment results, the approach can achieve a 92% accuracy rate [16].

Several studies combined ML algorithms, such as Decision Tree, Naive Bayes (NB), KNN, and SVM, to analyze the selection of performance features in breast cancer prediction. The model performs the training process using the Wisconsin Prognostic Breast Cancer (WPBC) dataset with 569 digital images. Among these methods, SVM has the best performance. The prediction performance of the decision support system is better than any individual ML model with High or Medium predictive confidence. Based on experimental results, SVM can achieve the accuracy of breast cancer prediction to 96%, which means it can provide robust assistance to doctors and patients [17].

Another study proposed employing three classifiers: NB, Bayesian Networks (BN), and Tree Augmented Naive Bayes (TAN) to analyze the predictive distribution of all classes using a comprehensive Bayesian methodology. First, the model conducts a training process with three datasets: breast cancer, Wisconsin, and breast tissue datasets. Next, they compared the prediction accuracy of the Bayesian approach with three other ML algorithms, namely KNN, SVM, and Decision Tree. BN achieves the best performance, with an accuracy of 97.281% [18].

Current research is exploring CNN to detect and classify images [19][20][21]. A paper introduces a Bilinear Convolutional Neural Network (BCNNs) approach for fine-grained categorization of histopathology images. The study employed bilinear pooling to combine various components without considering illness location into account, then compared several deep learning algorithms for fine-grained categorization to evaluate the model. Overall, the suggested methods are highly efficient, with accuracy rates of 99.24% in binary classification and 95.95% in fine-grained classification, respectively [12].

Another article combined the k-means Gaussian Mixture Model (GMM) method and CNN to classify breast cancer. There are three steps to create the proposed method: The first is to identify the region of interest (ROI). The second phase involves extracting ROI texture features and optimizing elements with an optimized feature selection method. The final stage used CNN to classify the disorder as benign or malignant. The model conducts a training process by Mammographic Image Analysis Society (MIAS) Dataset. The proposed method has an accuracy rate of 95.8% [22].

In the problem of classifying the histopathology features of breast cancer, a paper proposed an ensemble model. First, this model applied all individual models, including the well-trained VGG16 model, the well-tuned VGG16, the fully-trained VGG19, and the well-tuned VGG19 model. After they followed the ensemble strategy by taking probabilities, the fine-tuned ensembles VGG16 and VGG19 showed competitive classification performance, especially in the carcinoma class. They followed the ensemble strategy by taking possibilities. The VGG16 and VGG19 models achieved a sensitivity of 97.73% for
carcinoma grade and overall accuracy of 95.29%. In addition, he has an F1 score of 95.29%. The proposed method effectively classifies complex histological features of breast cancer [23].

Therefore, this paper proposes a model for breast cancer detection by introducing a new method using Faster R-CNN based on the feature dataset. To conduct our study, we apply a dataset of annotated images and original images of breast ultrasound. Wherewith this feature, the model can detect the location of breast cancer accurately.

3. Proposed Method

This section will formally define the research problem and some of the journal’s concepts.

A. Problem Definitions

This study uses Faster R-CNN to detect breast cancer based on annotated images and breast actual images. Fast R-CNN is a derivative of R-CNN. By introducing an RPN layer, this approach helps overcome the problem of slower R-CNN performance. The feature mapping obtained is entered into the RPN network to find prospective target regions, then input the mapping features and potential target areas through the ROI network. The Faster RCNN network accepts images of any size as input. Before entering Faster RCNN, we scale picture normalization, such as M×N. After that, we’ll be able to add a P×Q image size. If P×Q is higher than M×N, this image will be scaled, and if P×Q is smaller than M×N, the edge of the image will be filled with 0. Because of the Faster RCNN network’s structural characteristics, its loss function is similarly a multi-tasking loss, consisting of the target frame’s prediction loss and the target frame’s regression loss [24].

B. Proposed Method

A more sophisticated form of Fast R-CNN is Faster R-CNN. R-CNN has a derivative called Fast R-CNN. Region proposal generation occurs before the convolution layer. When dealing with huge photos, this step is said to cause slower performance. Faster R-CNN proposed addressing the performance issue by implementing the RPN layer and deleting the present region proposals generation. Faster R-CNN proposed addressing the performance issue. After executing feature extraction, the model calculates RPN [25].

The region proposal method, which provides bounding boxes or locations for probable objects in the image, is the first component of Faster R-CNN. Second, we use a feature generation stage to extract features from these objects, which is commonly done with the help of a CNN. The third layer is a classification layer that predicts which class this object belongs to. The fourth layer is a regression layer that determines the bounding box coordinates of the object. Fast R-CNN uses a CPU-based selective search algorithm that takes roughly 2 seconds per image and works on CPU computation for region proposals. The Faster R-CNN study addresses this by generating regional recommendations using the RPN. As a result, the Faster R-CNN reduces region proposal time and allows the region proposal stage to share layers with subsequent detection stages, resulting in a higher total feature representation [26]. Equation (1) In Faster RCNN, an image’s objective function is defined as:

\[
L([P_i]), [t_i] = \frac{1}{N_{cls}} \sum_i L_{cls} (P_i, P_i^*) + \lambda \sum_i P_i^* L_{reg} (t_i, t_i^*)
\]  

(1)
Equation (2) $P_i$ is the probability that the anchor will be predicted as a target.

$$ P_i^* = \begin{cases} 0 & \text{negative label} \\ 1 & \text{positive label} \end{cases} $$ (2)

Equation (3) is a vector that represents the predicted bounding box’s four-parameter coordinates. $t_i^*$ is the coordinate vector of the positive anchor in the ground truth bounding box.

$$ t_i = \{t_x, t_y, t_w, t_h\} $$ (3)

Equation (4) is a binary classification cross-entropy loss (target & non-target).

$$ L_{cls}(P_o, P_i^*) = -\log [P_i^*P_i + (1 - P_i^*)(1 - P_i)] $$

$$ L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) $$ is used to $L_{reg}(t_i, t_i^*)$

Equation (5) which is the regression loss, $R$ is the smooth function $L1$

$$ smoothL1(x) = \begin{cases} 0.5x^2 \times \frac{1}{\sigma^2} & |x| < \frac{1}{\sigma^2} \\ |x| - 0.5 & \text{otherwise} \end{cases} $$ (5)

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_i^*$</td>
<td>Probability that the anchor is predicted to be target.</td>
</tr>
<tr>
<td>$t_i^*$</td>
<td>The coordinate vector of the ground truth bounding box corresponding to the positive anchor.</td>
</tr>
<tr>
<td>$t_i$</td>
<td>The vector, which represents the four parameter coordinates of the predict bounding box.</td>
</tr>
<tr>
<td>$L_{cls}$</td>
<td>The cross-entropy loss of binary classification (target &amp; non-target)</td>
</tr>
<tr>
<td>$L_{reg}$</td>
<td>the regression loss</td>
</tr>
<tr>
<td>$R$</td>
<td>Smooth L1 function.</td>
</tr>
<tr>
<td>$L1$</td>
<td>Function.</td>
</tr>
</tbody>
</table>

Table 1. Mathematic notation of the Faster R-CNN

In the R-CNN process, each suggested region for each image requires feed-forward from CNN, even though the regions may overlap. Second, R-CNN runs three separate models: a feature extraction model, a classification model, and a regression model. Faster R-CNN is a new solution to deal with both problems. The overall picture, as well as the extraction stage, have both been improved. Therefore, there are two elements to make R-CNN Faster even better. First, unlike R-CNN, it integrates all models into a single network, including feature extraction, classification, and detection. Second, the number of times a regional CNN has to be run has been reduced to one per image. The term faster refers to the evolution of Faster R-CNN, which produces faster results [24].
4. Experimental Setup

A. Main Idea
The main goal of this paper is to use the Faster R-CNN algorithm to develop a model based on breast ultrasound images to detect breast cancer. For detecting and categorizing breast lesions, faster R-CNN comprises two key processes [27]. Faster R-CNN starts with an ultrasonic image as input and outputs a rectangle box around the desired item. Second, the RPN, trained with ground truth data to give Regional Proposals, passes the convolution feature map through it. Therefore, the feature map is sent into the RPN, which generates a set of predictive-score areas [25] [28].

B. Dataset
We collect a dataset of ultrasound images of breasts from kaggle.com to detect breast cancer in this experiment. Then we divide it into training and testing datasets. Learning models create using training datasets, whereas testing datasets to evaluate the models’ performance or accuracy. Here, the researcher divides it into 80% for training and 20% for testing for dataset distribution. Table 2 shows the details in terms of distribution of dataset used in the study, as follows:

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Training (80%)</td>
<td>840</td>
</tr>
<tr>
<td>Data Testing (20%)</td>
<td>210</td>
</tr>
<tr>
<td>Total</td>
<td>1,050</td>
</tr>
</tbody>
</table>

C. Data Pre-Processing
In this stage, we conduct pre-processing to process high-resolution photographs; because high-resolution photo processing takes a long time, it must reduce the image size. Then, we convert the image to grayscale. After that, we perform noise removal to find and remove unwanted noise from the digital image. We sort each sample price by magnitude. The sample median in the window is the middlemost value, which can be a filter output. Grouping photos into segments is necessary for recording changes to image attributes. Image analysis is performed pixel by pixel after segmentation, and each pixel is labeled based on whether the gray level pixel is greater or less than the threshold value. As a result, segmented image analysis becomes easier [29].

D. Detection Method
To conduct our study, we used Faster R-CNN to detect breast cancer. We collected a dataset containing original and annotated breast images. This study builds a detection model of breast cancer with a training and testing process. Before the pre-processing stage, we split the dataset into two classes of data: Annotate and image. The pre-processing is to simplify a detection model using Faster R-CNN. In this stage, we scale the image, remove unwanted noise from the digital image, then create a bounding box for the location of the breast cancer target, and then label it according to the image. We perform feature extraction to get vector values before feeding them to the training and testing process. After the pre-processing stage, we use the training dataset to train the model. We utilize the testing dataset to evaluate model performance using data validation in the testing phase of the process. Furthermore, a valid model was obtained and then tested using vector test data, which assessed the model's effectiveness in detecting breast cancer spots. Finally, after several stages, we obtain the Faster R-CNN that
can detect areas of breast cancer. We modified several settings to obtain the best performance detection to obtain the best accuracy value.

5. Result & Analysis

In the training process, this study employs original and annotated breast images. We also set epochs = 50 to train the model by adjusting various hyperparameters to get the best performance. A low error rate indicates that the model achieves good performance. Fig. 1 shows the loss vs. epoch score in this process.

Fig.1: Loss Vs Epochs Of Training Process

Fig.1 shows that at epoch 0, the loss result is 0.2743, while at epoch 49, the loss result is 0.0780. These results indicate that more epochs can produce in a tiny loss score. Based on the training result, our proposed model is good enough to help detect breast cancer with robust detection results.

In the testing process, we analyze 88 annotated data of breast cancer to test the model and get a red bounding box to mark the location of breast cancer. The testing process produces expected output and model output detection. Fig. 2 shows expected output and model output detection.

In Fig. 2, expected output indicates the actual detection of breast images. In contrast, the model output indicates the detection using our model. Based on the testing result, we can compare that the expected output just displays one bounding box and the model output can present more bounding boxes.

In the next phase of the testing process, we calculate actual detection and model
detection based on the number of breast cancer spots marked with a bounding box. The actual detection and model detection results are presented in graphical form. Fig. 3 shows the actual detection results and the detection results of our model.

Fig. 3: Actual vs Predicted of testing process

Fig. 3 shows the blue line that indicates the result of the actual test of real breast image detection, while the orange line indicates the result of model detection. Based on the testing result, our model can detect cancer more accurately than the real detection image approach.

In the next process, we calculate the ratio value of actual detection and model detection. Fig. 4 shows the actual vs. Predicted ratio

Fig. 4: Actual vs predicted ratio of testing process

Fig. 4 displays a graph of the actual detection ratio with our model's detection. The ratio graph shows the number of detection results from our model divided by actual detection results. Our model can harvest better detection than the actual detection in breast image detection issues. In the previous dataset, there were redundant and overlapping bounding boxes. To reduce this error, we implement Non-Maximum Suppressions (NMS). Fig. 5 shows NMS model output of breast cancer detection.
Fig. 5. shows detection using NMS to reduce detected bounding boxes in the testing process. At the most basic level, most object detectors perform some form of windowing. Object detectors generate thousands of windows (anchors) of various sizes. After the detector removing many bounding boxes, it is necessary to filter the best to avoid window overlapping. After training and testing process, our proposed model can achieve better detection results than conventional breast cancer detection techniques. Therefore, we can conclude that our proposed model can detect more objects.

6. Conclusion

Detection breast ultrasound images is a major difficulty in breast cancer detection. To solve this challenge, the current papers recommend using traditional machine learning. However, it's expensive and time-consuming in manual feature engineering. Therefore, to enhance the detection model's performance with automatic feature engineering, this research recommends using Faster R-CNN algorithm to develop a breast detection model.

Based on the experimental result, the Faster R-CNN can achieve higher accuracy and tiny loss during the training process. In the training process, the study set some hyperparameters Epoch = 50. The training process produces a loss of 0.0780, region box loss of 0.0471, abjectness loss of 0.0005, and RPN box loss of 0.0014. In the testing process, the study produces more bounding boxes than the conventional technique. Therefore, the model can be a promising solution to deal with breast cancer detection challenges accurately and in real-time.

Future research can adopt another algorithm to improve this model, such as GAN architecture. GAN can produce good image capabilities and provide an accurate solution for medical image analysis. The use dynamic neural network can expect to produce higher quality accuracy with additional features that can develop.

Acknowledgment
This paper is conducted in the Department of Informatics, Respati University of Yogyakarta, Indonesia.
References


