CHAPTER 1

INTRODUCTION

This chapter includes the following subtopics, namely: (1) Rationale; (2) Theoretical Framework; (3) Conceptual Framework/Paradigm; (4) Statement of the problem; (5) Hypothesis (Optional); (6) Assumption (Optional); (7) Scope and Delimitation; and (8) Importance of the study.

1.1 Rationale

Breast cancer is prevalent among women and ranks as a leading cause of female mortality worldwide [1]. The World Health Organization (WHO) [2] reported that in 2020, approximately 2.3 million women were diagnosed with breast cancer, resulting in 685,000 deaths globally. By the end of 2020, 7.8 million women were living with a breast cancer diagnosis from the past five years. Given the substantial incidence of breast cancer, it is imperative to prioritize early detection through screening to mitigate the prevalence of the disease. Statistics show that around 40% reduction in deaths from breast cancer can be achieved if early detection is carried out [3]. This explains why regular breast cancer screening is essential [4].

One of the examination procedures for detecting breast cancer is using the ultrasound screening method. Ultrasound screening is a common technique for diagnosing breast cancer, as it can differentiate between solid masses and fluid-filled areas of the breast. Also, it can identify irregular shapes and additional cancer blood vessels [5]. That is why ultrasound has recently become a significant concern for dense breasts [6]. However, medical personnel usually need a lot of time and energy to read or analyze ultrasound results. In addition to the increasing volume of ultrasound in hospitals, the classification process is also tricky due to spot noise, complex textures, and subjective judgment of medical personnel [7, 8]. Currently, many methods can be used to assist medical personnel in classifying breast cancer from ultrasound images [9–11].

Artificial Intelligence plays an essential role in breast cancer classification. Deep learning has provided better results in image processing with the ability to automatically learn complex features from images [12]. One of the uses that can be applied in image processing methods is a field of computer vision that can help classify types of breast cancer into specific categories. Shehab et al. [13], an analysis of 200 publications spanning from 2000 to 2022 focused on the utilization of machine learning in medical application. The research explains that machine learning can examine a lot of data, find exciting relationships, and perform pattern recognition, ultimately enhancing the effectiveness and precision of diagnostic systems across various diseases, including cancer detection. Sebastian and Peter

[14] discusses AI in cancer research; it explains that humans have limitations in analyzing large amounts of data in a short time; the integration of AI into cancer research overcomes the mistakes of medical experts in failing to diagnose and cure cancer, and the results are pretty encouraging. AI powered systems have the potential to assist pathologists in making more precise cancer diagnoses, thereby lowering error rates. As a result, the integration of AI, particularly through machine learning, in automation is significantly enhancing USG based breast screening processes.

In the context of ultrasound classification of breast cancer, several studies have demonstrated significant efficacy. Xiao et al. [15] compared transfer learning, CNN3, and conventional machine learning. These findings show that the transfer learning model using inceptionV3 provides superior results compared to other models. This underscores the challenges in developing CNN architectures to achieve optimal performance, which depends on several factors, such as the number of layers in the network, the number of filters in a layer, the convolutional size, the depth of the network, and the large number of available datasets [12]. Mukhlif et al. ; Wu et al. [7, 8] also applied transfer learning to ultrasound breast cancer classification. The results show that transfer learning can significantly improve predictions, producing more accurate results using relatively small data sets. Integrating AI, particularly through machine learning and deep learning, has significantly improved the breast cancer diagnostic process. Using these techniques in medical image processing, such as ultrasound, helps overcome human limitations in analyzing large and complex data, thereby providing more accurate and efficient results in breast cancer diagnosis.

However, despite significant progress, the research must improve accuracy and efficiency. The use of more complex and deep models, such as VGG-16, InceptionV3, and ResNet-50, requires large computational resources [8, 16, 17], this approach sometimes fails to improve network efficiency in size and speed. On the other hand, [9, 18] suggest further research to use lightweight models such as MobileNet or MobileNetV2, which can work well if set up correctly because transfer learning models are still widely used with computing high on medical images. In addition, although deep learning systems can provide correct diagnosis results, doctors are often unable to understand the basis or reasoning behind the decisions produced by the system, which gives rise to a new challenge called the ”Black Box” problem [19–21]. This lack of clarity hinders acceptance and trust in this technology among medical professionals. Therefore, the system must also explain the reasons for its development. Fujioka et al. [19] recommend the use of Grad-CAM.

Based on the abovementioned problems, this research aims to overcome the gaps by proposing a new model using MobileNetV2 combined with the Convolutional Block Attention Module (CBAM). This combination is expected to increase accuracy in diagnosis and speed up the analysis process. To strengthen explanations in breast cancer classification and overcome the “Black Box” problem, this research also applies Gradient Weighted Class Activation Mapping (Grad-CAM). Grad-CAM visualizes areas of a medical image that the model focuses on when making decisions. Thus, this technique improves diagnostic accuracy and provides greater transparency in the decision-making process. This allows medical personnel to understand the basis of the results provided by the model, thereby increasing trust and acceptance of this technology among medical professionals.

Therefore, it is very important to use image processing techniques to classify breast cancer on ultrasound. One image processing method currently popular and proven to provide high performance is the MobileNetV2 model [22] as well as an attention module that utilizes CBAM [23]. This module can identify the most relevant regions and significantly increases accuracy. Using breast cancer ultrasound images from two databases, a strategic approach has been used to assess the overall performance of the general model. Given the limited size of the dataset, even after combining images from both databases, applying transfer learning with established pre-training is considered advantageous. This approach allows the classifier to leverage existing knowledge and avoids the need to train models from scratch. Grad-CAM is also used to generate heat maps to visualize features by calculating classification gradients and feature maps in the final convolution to achieve a focus on differentiating normal, benign, and malignant cancers in ultrasound breast cancer.

1.2 Theoretical Framework

1.2.1 Breast Cancer and Its Diagnosis

Breast cancer remains a critical health issue globally, impacting millions of women and contributing significantly to female mortality. According to [1], and corroborated by the World Health Organization [2], the incidence and mortality rates of breast cancer are alarming. Early detection is crucial for improving survival rates, as evidenced by studies indicating that early screening can reduce breast cancer mortality by up to 40% [3]. Regular breast cancer screening is essential, as emphasized by [4].

1.2.2 Ultrasound Screening in Breast Cancer

Ultrasound screening is a widely used diagnostic tool for breast cancer due to its ability to distinguish between solid masses and cystic structures, and to detect irregular shapes and blood flow indicative of malignancies [5]. Ultrasound is particularly significant for women with dense breast tissue [6]. Despite its benefits, the analysis of ultrasound images is labor-intensive and prone to variability due to factors like noise, complex textures, and the subjective judgment of medical personnel [7, 8].

1.2.3 The Role of Artificial Intelligence in Medical Imaging

Artificial Intelligence (AI), especially through deep learning, has revolutionized image processing by automatically learning complex features from images [12]. AI can analyze large datasets, identify patterns, and enhance diagnostic accuracy and efficiency [13, 14]. In breast cancer diagnostics, AI-powered systems help mitigate human errors and improve the precision of diagnoses.

1.2.4 Deep Learning Models in Breast Cancer Classification

Several studies have demonstrated the efficacy of deep learning models in breast cancer classification using ultrasound images. [15] found that transfer learning models, particularly those using InceptionV3, outperformed traditional machine learning methods. These models can significantly improve predictive accuracy even with small datasets [7, 8]. However, deep learning models like VGG-16, InceptionV3, and ResNet-50 require substantial computational resources [16, 17], posing challenges in terms of efficiency and speed.

1.2.5 Lightweight Models and the Black Box Problem

To address these challenges, research suggests using lightweight models such as MobileNet or MobileNetV2, which are efficient and effective when properly configured [9, 18]. Additionally, the black box problem, where the decision-making process of AI models is not transparent, poses a significant barrier to the acceptance of AI in medical diagnostics [19, 20].

1.2.6 Proposed Model: MobileNetV2 with CBAM and Grad-CAM

This research aims to address these issues by proposing a model combining MobileNetV2 with the Convolutional Block Attention Module (CBAM) to enhance diagnostic accuracy and efficiency. MobileNetV2 is known for its high performance in image processing with minimal computational requirements [22]. CBAM enhances model accuracy by identifying the most relevant regions of the image [23]. Additionally, Grad-CAM will be used to generate heat maps for visualizing important features, thereby improving the model's explainability and acceptance among medical professionals.

1.3 Conceptual Framework/Paradigm

1.3.1 Identification of Variables

In this research, several key variables are identified that are crucial for the effective classification of breast cancer using ultrasound images:

Independent Variables:

1. Ultrasound Images: The raw input data consisting of breast ultrasound images.
2. MobileNetV2: A lightweight convolutional neural network model used for feature extraction.
3. Convolutional Block Attention Module (CBAM): An attention mechanism to enhance feature representation.
4. Grad-CAM: A visualization technique to generate heat maps for model interpretability.

Dependent Variable:

Breast Cancer Classification: The output of the model, categorizing the ultrasound images into normal, benign, or malignant.

Moderating Variables:

1. Image Quality: The quality and resolution of the ultrasound images, which can affect model performance.
2. Computational Resources: The hardware and software resources available for model training and inference.
3. Medical Expertise: The level of expertise of medical personnel in interpreting the results.

1.3.2 Discussion of Relationships

Ultrasound Images:

Serve as the primary data input for the classification model. The quality and resolution of these images are crucial for accurate feature extraction.

MobileNetV2:

1. Acts as the base model for feature extraction due to its efficiency and performance.
2. Processes the ultrasound images to extract relevant features while maintaining computational efficiency.

Convolutional Block Attention Module (CBAM):

1. Enhances the extracted features by focusing on the most relevant regions of the images.
2. Improves the model's accuracy by enabling it to attend to important areas in the ultrasound images.

Grad-CAM:

1. Provides a mechanism to visualize the important features identified by the model.
2. Generates heat maps that highlight areas of the ultrasound images contributing to the classification, improving explainability and trust in the model's decisions.

Breast Cancer Classification:

1. The final output of the model, categorizing the ultrasound images into normal, benign, or malignant.
2. The accuracy and reliability of this classification depend on the effectiveness of the preceding stages.

Moderating Variables:

1. Image Quality: High-quality images lead to better feature extraction and, consequently, more accurate classifications.
2. Computational Resources: Adequate resources ensure efficient model training and faster inference times.
3. Medical Expertise: Expertise in interpreting the results can validate the model's outputs and provide insights into areas needing improvement.

The conceptual framework outlines the relationships between the various elements and variables involved in the research. By understanding these relationships, the proposed model aims to enhance the accuracy, efficiency, and interpretability of breast cancer classification using ultrasound images. The integration of MobileNetV2, CBAM, transfer learning, and Grad-CAM is expected to provide a robust solution that addresses current limitations and improves diagnostic outcomes.

1.4 Statement of the Problem

Previous studies on ultrasonographic breast cancer classification have made comparisons of deep neural network transfers [15]. In this research, a comparison of transfer learning, CNN3 and conventional machine learning was carried out. Based on the results obtained, the transfer learning model with InceptionV3 achieves the best results with an accuracy of 85.13% outperforming other models. This explains that there are challenges in developing the CNN architecture to achieve good results. It depends on the amount of medical image data which is known which is a difficult task to obtain and the architecture such as the number of layers in the network, how many filters in a layer, how many convolutional sizes, network depth, learning rate, etc. [12]. Usually, choices are made only by manual experimentation, which of course requires a certain level of knowledge and expertise in deep learning and trial and error can produce disappointing results for beginners [12]. Because of this, it is inefficient to try all possibilities manually to find the best model.

Many research have been conducted on ultrasound classification of breast cancer using transfer learning [7, 8, 15]. The problem of this research is using a pre-trained model with large parameters so that it has high computing power such as ResNet-50, VGG-16, InceptionV3 and others. The general trend is to create deeper and more complex networks to achieve higher accuracy, this does not necessarily make networks more efficient in terms of size and speed. As in the main reference paper, namely [8], an ultrasound classification of breast cancer was made using the VGG-16 model which is expected to help the clinical application of breast cancer diagnosis widely. However, it should be able to use a lightweight architecture model such as MobileNetV2 [22] which is specifically designed for mobile environments. Therefore, in this research we wanted to optimize limited resources by reducing the number of operations and memory while maintaining the same accuracy. So, it needs to be examined and focused on ease of development to this mobile terminal. The results of this research achieved an accuracy of 91%, in two datasets namely [24, 25] with benign and malignant categories. However, the 3 categories (normal, benign and malignant) can actually be used with accuracy which can be improved by suggestions using the CBAM attention module.

One of the important properties of the human visual system is that a person does not try to process the whole scene at once. Instead, they focused specifically on protruding parts to better capture the visual structure. The weakness of the current ultrasound breast cancer classification research is that it only focuses on using convolutional models and no one has yet implanted the CBAM attention module [23] to solve the problem of ultrasound breast cancer classification. Whereas, CBAM is a lightweight module and can be integrated into the CNN architectural model, which can help the model improve its ability to learn and focus on important features in images that allow this information to be used for more accurate predictions. Comparative evaluation was carried out with and without CBAM.

The results showed that CBAM not only significantly improves accuracy, but also increases efficiency, because CBAM is quite small both in terms of parameters and computation, so it has the potential to be applied to low-end devices [23]. Indeed, someone has completed research on breast cancer classification with an architectural model by embedding CBAM, but the images used are mammogram images [26] and histopathology [27].

Transfer learning has been shown to improve performance in cases of breast cancer classification [7, 15, 28] as well as in cases such as driver behavior detection [29] and lotus species classification [30]. However, its weakness is explanatory ability which has to do with understanding the reasons behind model decisions [19–21]. Therefore, we need a method to improve the ability to explain in image recognition. In this research, due to its excellent visual support, Grad-CAM was used to determine the localization ability of the target area at the final convolution[31]. This will help in providing an interpretation of why the model made certain decisions. Based on the problem descriptions from previous studies, this research will classify ultrasound breast cancer using the MobileNetV2 transfer learning method, which is specifically designed for cellular environments and also embedded with a CBAM layer to increase representational power. Grad-CAM was also used to visualize the targeted region in the final convolution. With this the authors introduce the MobileNetV2-CBAM model in the classification of breast cancer on USG images. The research questions to be answered in this reseach are:

1. Does the MobileNetV2 pre-trained model with CBAM attention module provide good accuracy and low computation compared to complex models (VGG-16, VGG-19, InceptionV3, ResNet-50) in ultrasound breast cancer classification?
2. How does the use of Grad-CAM visual influence the decision interpretation of image recognition models?

1.5 Objective

The objectives of this thesis research are:

1. Improve the accuracy of ultrasound breast cancer classification by embedding the Convolutional Block Attention Module (CBAM) in the MobileNetV2 architecture, thus enhancing the model's ability to identify important regions and improve classification performance.
2. Reduced computation using Transfer Learning MobileNetV2 with CBAM, a lightweight and efficient architecture that allows the model to be deployed on resource-constrained devices, compared to more complex models (VGG-16, VGG-19, InceptionV3, and ResNet50).

3. Improve explainability by implementing Gradient-weighted Class Activation Mapping (Grad-CAM) to provide a better visual interpretation of how the model makes classification decisions.

1.6 Hypotheses

The MobileNetV2 pre-trained model has a small number of parameters which is 3.4 million[12], which is compared to the number of parameters of other models such as VGG-16 : 138 million[25] , VGG-19 : 144 million [25], InceptionV3: 27 million [26] and ResNet-50: 25,6 milion [27], of course the MobileNetV2 model is the right solution for devices with limited resources. Several studies[22] [28], state that MobileNetV2 has the best accuracy compared to other models. This proves that MobileNetV2 offers good performance while maintaining high execution speed with good accuracy. In order to improve the model in identifying important areas and improve classification performance, the CBAM attention module is used. Studies have shown[13][29] that embedding a CBAM in a pre-trained model does not significantly increase the model size, but significantly increases the accuracy of the model making it more suitable for deployment to cellular terminals. Grad-CAM will improve explainability to provide a visual interpretation of how the model makes classification decisions, so that later you can see the Heatmap of the proposed method.

Therefore, the hypothesis of this research is that Using the pre-trained MobileNetV2 model with the embedded attention module CBAM provides better model performance and accuracy compared to the MobileNetV2 model without the CBAM, VGG-16 and other models. Furthermore, the incorporation of Grad-CAM will enhance the model's explainability, allowing users to understand and visualize the critical regions identified by the model during classification. This increased transparency and explanability will make the model more reliable and trustworthy.

1.7 Assumption

This research is based on the following assumptions:

1. Early Detection Reduces Mortality:

   Early detection of breast cancer significantly reduces mortality rates.

2. Effectiveness of Ultrasound:

   Ultrasound imaging is effective for breast cancer screening, especially in women with dense breast tissue.

3. Improved Accuracy with Deep Learning:

   Deep learning models, such as MobileNetV2 and CBAM, improve the accuracy of breast cancer classification from ultrasound images.

4. Benefit of Transfer Learning:

   Transfer learning with pre-trained models enhances model performance with limited datasets.

5. Visualization Enhances Interpretability:

   Grad-CAM visualization improves the interpretability of the model's decisions.

6. Importance of Image Quality:

   High-quality ultrasound images lead to better classification results.

7. Impact of Computational Resources:

   Adequate computational resources are essential for efficient model training and inference.

8. Role of Medical Expertise:

   Medical professionals' expertise is crucial for validating and applying the model's results.

1.8 Scope and Delimitation

1.8.1 Scope

Principal Variables:

1. Independent Variables: Ultrasound images, MobileNetV2, CBAM (Convolutional Block Attention Module), and Grad-CAM.
2. Dependent Variable: Breast cancer classification (normal, benign, malignant).
3. Moderating Variables: Image quality, computational resources, and medical expertise.

Locale:

This research utilizes ultrasound image datasets from publicly available databases. No specific geographical region is targeted for data collection beyond these public datasets.

Timeframe:

The research is conducted over a period of one year. This includes data collection, preprocessing, model development, training, evaluation, and analysis.

Justification:

1. Selection of Variables: The choice of independent variables such as MobileNetV2 and CBAM is justified by their proven effectiveness in image classification tasks. Transfer learning is employed to improve performance on limited datasets, while Grad-CAM is used to enhance model explainability.
2. Locale: The use of publicly available databases ensures a diverse and comprehensive dataset, which is essential for developing a robust model. It also allows the study to be reproducible by other researchers.
3. Timeframe: One year period is deemed sufficient to carry out the various phases of the research, from initial data handling to final analysis, ensuring thorough and detailed investigation within a reasonable duration.

1.8.2 Delimitation

Exclusion of Other Imaging Modalities:

The study does not consider other breast cancer detection methods such as mammography, MRI, or CT scans, focusing solely on ultrasound images.

Geographical Limitation:

The research is confined to publicly available ultrasound image datasets and does not incorporate data from local hospitals or clinics, potentially affecting the generalizability of findings.

Model Complexity:

The study is limited to using MobileNetV2 and CBAM due to constraints on computational resources.

Patient Demographics:

The study does not take into account patient-specific variables such as age, ethnicity, or medical history, which could impact the classification outcomes.

Focus on Explainability:

While Grad-CAM is used to improve the explainability of the model, other explainability techniques are not explored or compared within this study.

1.9 Significance of the Study

This study aims to contribute to the field of medical imaging and breast cancer diagnosis by proposing an advanced model for the classification of breast cancer using ultrasound images. The significance of this study is multifaceted, impacting various stakeholders and advancing the existing body of knowledge.

1.9.1 Contributions as New Knowledge

Enhanced Diagnostic Accuracy:

1. By integrating MobileNetV2 and CBAM, the study provides a new approach to feature extraction and attention mechanisms, resulting in improved accuracy in classifying breast cancer from ultrasound images.
2. The use of transfer learning with pre-trained models helps in achieving high performance even with limited datasets, showcasing a practical solution for real-world applications.

Improved Model Explainability :

The application of Grad-CAM for visualizing important features in the images enhances the explainability of the model's decisions, addressing the black-box problem commonly associated with deep learning models.

Efficiency in Medical Imaging:

The proposed method offers a lightweight and efficient solution for breast cancer classification, which can be implemented in resource-constrained settings without sacrificing performance.

1.9.2 Usefulness to Specific Groups

Medical Professionals:

1. Radiologists and oncologists can benefit from the improved accuracy and explainability of the classification model, aiding in more precise and reliable diagnosis of breast cancer.
2. The visualization provided by Grad-CAM can help medical professionals understand the model's decision-making process, thereby increasing their trust in automated diagnostic tools.

Researchers and Academics:

1. This study provides a comprehensive framework that can be further explored and refined by researchers in the field of medical imaging and artificial intelligence.

2. The findings contribute to the growing body of literature on the use of deep learning in healthcare, offering insights and methodologies that can be applied to other medical imaging tasks.

Healthcare Institutions:

1. Hospitals and clinics can implement the proposed model to enhance their diagnostic capabilities, particularly in regions with limited access to advanced medical imaging technologies.

2. The efficiency and accuracy of the model can lead to better patient outcomes, reducing the burden on healthcare systems and improving overall public health.

Patients:

1. Early and accurate detection of breast cancer can significantly improve treatment outcomes and survival rates for patients.

2. By providing a reliable diagnostic tool, the study contributes to more timely and effective management of breast cancer, ultimately benefiting the patients who rely on accurate diagnoses for their treatment plans.