

Plagiarism - Report

Originality Assessment

17%



Overall Similarity

Date: Apr 17, 2024

Matches: 573 / 3423 words

Sources: 25

Remarks: Low similarity detected, check with your supervisor if changes are required.

Verify Report:

2 Breast Cancer Detection Using Texture Analysis and Convolutional Neural Network

Ashutosh Gupta

1Student, Dept. of CSE

Shri Sant Gajanan Maharaj

College of Engineering, Shegaon

gashutosh805@gmail.com

Surbhi Gorla

2Student, Dept. of CSE

7 Shri Sant Gajanan Maharaj

College of Engineering, Shegaon

surbhigoria168@gmail.com

Pallavi Awasare

3Student, Dept. of CSE

Shri Sant Gajanan Maharaj

College of Engineering, Shegaon

awasarepallavigajanan@gmail.com

Abstract-- Breast cancer is a big problem for women all over the world and it can make them very sick. Researchers have worked hard to improve how we diagnose and detect this disease accurately. It remains 2 one of the most life-threatening illnesses, affecting about one in eight women. The unclear causes make it challenging to manage, and prevention is difficult. So, early detection becomes crucial. This paper wants to explain a clear way to find breast cancer early by using computer tools to analyze images. It will explain the steps involved, such as image enhancement, segmentation, and feature extraction, utilizing a Convolutional Neural Network (CNN).

Keywords: Diagnose, Prevention, Crucial, Enhancement, Segmentation

1. 2 INTRODUCTION

Breast cancer affects a large number of women worldwide. While it can be found through a breast exam, it's often hard to detect, especially the areas you can't feel. But these hidden areas show up clearly on mammograms or ultrasound scans.

Obtaining specialized images of the breast, known as mammograms, is considered the most effective method for detecting breast cancer at its earliest stages. However, reading these scans can be tricky. So, doctors use special techniques to help spot tumors. Cancer detection is tough because each case is unique. One person's cancer might react differently to treatment than another's. Also, using just one method to find breast cancer might not work well because every cancer and breast is different. If someone's had breast surgery, their mammogram might look different too. For the past two decades, breast cancer [3] has been a prominent area of research, resulting in significant advancements in detection methods. However, diagnosing and treating the disease remains challenging and costly. Finding masses automatically is still hard because each cancer is unique and needs its own treatment. We still have a lot to learn about breast cancer.[1]

Design and Implementation

This project work in both front end and backend. In this web project of breast cancer detection, the front-end typically includes everything that the user interacts with directly. This encompasses the [8] user interface, design, and user experience components. This project is about making the website look good and easy to use. We want users to be able to log in, sign up, and see their results without any trouble. To do this, we write special code that works in their web browser.

Fig. 1.1 Homepage

We use things like HTML, CSS, and JavaScript to make it happen. These technologies help [22] to create interactive elements, validate user input on forms, and handle other client-side interactions. The front-end of a web project focused on creating an engaging, visually appealing, and user-friendly [8] interface through which users can interact with the system and receive the results of the

breast cancer detection process. The website includes a login page for authorized access, a signup/registration page for new users, and information on precautions for breast cancer detection. These features aim **6** to provide a secure and user-friendly experience while also promoting awareness and proactive health management.

Fig. 1.2 Registration Page

While the back-end handles the processing and generation of results, the front-end is responsible for sending **data to the** back-end and displaying the results **to the user**. This involves integrating the front-end with the back-end through APIs **11** (Application Programming Interfaces) or other means of communication. In this project it ensure that user input is sent to the back-end **6** for processing, and the results are displayed back **to the user** in the front-end interface. **In this case**, this process is a breast cancer detection algorithm or model. This model examines the provided data and generates a result, indicating whether there are **3** signs of breast cancer or not.

Fig. 1.3 Model Design

2. LITERATURE REVIEW

Zahra Abdolali Kazemi and her team's research revealed that breast cancer contributes to 19 percent of cancer-related deaths and affects 24 percent of all cancer cases in European countries. It's surprising that many **3** women aged 40 to 49 die from breast cancer. Finding **breast cancer early** is very important, and computers help by finding any strange signs at the beginning. This chapter will explore various techniques and provide a qualitative comparison between them. The study evaluates two methods for presenting mammography images: displaying them simultaneously or alternately on the same screen. Image processing algorithms, including **10** Support Vector Machine (SVM),

Convolutional Neural Network (CNN), K-nearest neighbors (KNN), and Genetic Algorithm (GA), are employed. The performance of these algorithms will be thoroughly discussed. ¹⁹ The training process involves providing features related to different classes and updating parameters accordingly. Then, unlabeled data are classified based on this training. Segmentation, which simplifies or modifies image views for easier analysis, entails labeling each pixel in the image, resulting in segments covering the entire image. "Doctors can spot cancer cells and make diagnoses by looking at these segmented images. Adding different tools to MATLAB helps with tasks like practicing, teaching, and sorting through information.[1]

Saif Ali and his team describe cancer, also known as malignancy, as an irregular proliferation of cells. There are more than 100 types of cancer, ranging from skin cancer, colon cancer, prostate cancer, breast cancer, to lymphoma. Symptoms of lung cancer can vary depending on the specific type. ¹⁵ According to the American Cancer Society, it is projected that there will be 1,806,950 new cases of cancer in the United States this year, leading to 606,520 deaths. Cancer therapy may include surgery and/or chemotherapy. ¹² Cancer ranks as the foremost cause of death worldwide, categorized into malignant and benign types. Early detection proves pivotal for successful cancer management, prompting exploration into various detection methodologies. Manual identification proves labor-intensive and unreliable, leading to investigations into computer-aided detection. This involves image manipulation to isolate characteristics and classification strategies to discern cancer type and stage. The document explores various algorithms, such as SVM, KNN, DT, etc., for categorizing ¹² different cancer types. Moreover, it presents a comparative analysis of previous research endeavors.[2]

Mutiullah and his team discuss how lung cancer has become a major worry for people worldwide in recent years. This has led many countries to provide funds and seek help from scientists to tackle the disease. Researchers have suggested different ideas and faced challenges in developing computer systems to detect lung cancer early. They've also shared important facts about lung cancer. Computer Vision (CV) technology is key in the fight against lung cancer. Image processing plays a crucial role in

computer vision, especially in improving ¹⁷ the performance of medical diagnostic machines used in medical imaging. This involves implementing various technical procedures. These steps are vital for achieving accuracy comparable ² to that of other authors who use specific algorithms or techniques. This article highlights the typical steps that many authors follow in preparing, segmenting, and classifying methods to detect areas of lung cancer. If there is uncertainty in the preprocessing and segmentation steps, it can negatively impact the classification process. Let's briefly discuss these factors to help new researchers understand the situation and determine the best direction to proceed.[3]

Yousif M.Y Abdallah and colleagues found that improving mammography images is a strong way to categorize breast tissues and identify problems. Using special computer programs, mammographers can make the pictures clearer. In their study, they used methods like making colors clearer, reducing fuzziness, checking textures, and cutting up the pictures into smaller parts. They kept the pictures stored in high quality to keep them looking good. These methods aimed to make the pictures brighter and clearer while getting rid of any fuzzy spots. How well these methods worked depended on the type of tissue and how the background of the breast looked. They also looked at how fast the computer could do all this. They ³ found that the methods improved the detection of breast problems by about $96.3\% \pm 8.5\%$ ($p < 0.05$). So, by using these techniques, they showed that they could make it easier to spot problems in the breast images.[4]

Prannoy Giri and his team studied breast cancer, a big problem that causes many deaths among women. They did a lot of research to find out how to detect this deadly disease early. ¹¹ "At some point in their lives, many women encounter breast cancer. Because we're not sure what causes breast cancer, it's hard to prevent. Finding breast cancer early is really important. Using special computer programs to look at mammograms is a good way to find breast cancer. If we can find it early, we can save lives. Finding lumps or tiny ⁵ calcium deposits in the breast is a sign that someone might have breast cancer. These things can help us find cancer when it's just starting. The researchers used pictures from a big database of mammograms that's used all over the world for cancer research. They looked closely at the pictures to ¹⁴ find out what textures show up when someone has cancer. Then,

they used a special method to see which textures meant someone **might have cancer**. This helped them find cancerous lumps more accurately. Finally, they used a computer program to analyze the textures and figure out what patterns mean someone has cancer in a mammogram.[5]

Arpita Joshi and Dr. Ashish Mehta explored various methods including KNN, **SVM, random forest, and** decision tree for classifying results. They looked at information about breast cancer from a dataset called **the Wisconsin Breast Cancer dataset**, which they found in a place called the UCI repository. Their simulation results showed that KNN performed the best, followed by SVM, random forest, and decision tree classifiers.[6]

David A. Omon-diagbe, Shanmugam Veeramani, and Amandeep S. Sidhu investigated the effectiveness of **different machine learning techniques, including support vector machines**, artificial neural networks, and Naive Bayes, utilizing **the Wisconsin Diagnostic Breast Cancer (WDBC) Dataset**. They integrated these methods with feature selection and extraction techniques to identify the most optimal approach. Their simulations revealed that despite its longer computation time, SVM-LDA emerged as the preferred **method for breast cancer diagnosis** among the tested techniques.[7]

Kalyani Wadkar, Prashant Pathak, and Nikhil Wagh conducted a detailed comparison between **artificial neural networks** (ANN) and **support vector machines (SVM)**. They also incorporated different classifiers such as **convolutional neural networks** (CNN), K-nearest neighbors (KNN) to improve dataset processing. After analyzing **experimental results and performance metrics**, they **concluded that ANN** is more effective compared to SVM. [8]

Anji Reddy Vaka, Badal Soni, and Sudheer Reddy K. devised a novel technique **for breast cancer detection using machine learning** methods such as the Naive **Bayes Classifier, SVM classifier, and Bidirectional Recurrent Neural Networks** (HABiRNN). They compared these methods with their proposed approach, which utilized a **Deep Neural Network with Support Value**. The simulation

outcomes demonstrated that the DNN algorithm outperformed others in terms of both efficiency and image quality. This underscores the significance of advanced medical systems that prioritize these factors.[9]

Monica Tiwari, Rashi Bharuka, Praditi Shah, and Reena Lokare devised a novel approach to detect breast cancer. They employed various computer methods including logistic regression, random forest, K-Nearest Neighbor, Decision Tree, Support Vector Machine, and Native Bayesian Classifier.

Additionally, they utilized deep learning methods such as 2 artificial neural networks, convolutional neural networks, and recurrent neural networks. They compared these methods 3 and found that the CNN model was 97.3 percent accurate, and the ANN model was even better at 99.3 percent accuracy, beating the other computer methods.[10]

3. METHODOLOGY

Think of a 1 Convolutional Neural Network (CNN) as a smart tool that helps doctors find signs of breast cancer. It learns from lots of breast pictures to spot possible cancer signs. When a CNN sees a new picture, it uses what it knows to figure out if there might be cancer in it. This helps doctors see areas in the pictures that 14 might have cancer. So, it's like having extra eyes to find breast cancer early."

- Input Layer: This layer takes in the picture of the breast tumor, represented as tiny dots called pixels. Each pixel carries information about the intensity and color of the image.

- Convolutional Layer: It looks at different 1 parts of the picture and identifies important features or patterns that help determine whether the tumor is malignant or benign. These features could include edges, textures, or shapes within the image.

- Pooling Layer: 9 After the convolutional layer identifies important details, the pooling layer comes

in to condense the information. **1 It reduces the size of the** picture while retaining the essential features, much like focusing a camera to capture the main subject while blurring the background.

□ **Flatten Layer:** The flattened layer takes the condensed information from **the pooling layer** and converts it into a single-dimensional list. This restructuring helps the computer better understand and process **the features extracted from the image.**

□ **Fully Connected Layer:** In this layer, the network **combines all the** extracted features to make a decision about the nature **of the tumor.** It analyzes the relationships between different features and learns **to classify the tumor as** either malignant or benign **based on the** patterns observed.

□ **Output Layer:** This layer provides **the final output of the network,** indicating the likelihood that **the tumor is** malignant or benign. It gives a probability score, with higher values suggesting a higher likelihood of malignancy.

Fig. 3.1. Block Diagram of CNN

Each layer in the CNN plays a specific role in processing the image data and extracting relevant information for accurate classification. Through the iterative process of training on large datasets of labeled images, the CNN learns to optimize its parameters to effectively distinguish between **2 different types of breast** tumors.

By combining the strengths of each layer, CNNs have shown remarkable success in automating the analysis of medical images and assisting healthcare professionals in making informed decisions about patient care.

Fig. 3.2. Architecture **2** of Breast Cancer Detection

□ Image acquisition **for breast cancer detection using Convolutional Neural Networks (CNNs)**

typically involves obtaining mammogram images through specialized imaging equipment.

4 During a mammogram, the breast gets squeezed between two plates, and X-rays take pictures of the breast tissue. These images show the inside of the breast clearly, helping radiologists spot any unusual things like tumors or masses.

□ Preprocessing **2 for breast cancer detection using** CNNs involves preparing the images before feeding them into the computer program. This includes cleaning up any unwanted elements, adjusting brightness and contrast, standardizing image sizes, and generating more data through techniques like flipping or rotating images. These steps ensure that the images are in a suitable format for the program

to analyze and detect **3 signs of breast cancer** accurately.

□ Segmentation **13 for breast cancer** detection with CNNs involves outlining **the breast area** in images to focus on important regions. The program identifies the breast, draws boundaries around it, and then highlights any abnormalities within that area. This focused approach helps in accurately identifying potential **4 signs of breast cancer for** further examination.

□ The CNN program looks at the images **5 of breast tissue** and picks out important features, like textures or shapes, **that could be signs of cancer**. It does this by analyzing different parts **10 of the images and** finding patterns that are common in cancerous tissue. Once the program has extracted these features, it uses them to make a decision about whether cancer is present or not. It compares the features it **2 found in the** images to patterns it learned during training, and then decides if the images show signs of cancer or not.

i. Positive result indicates that the CNN model has identified signs or features in mammogram images that suggest the presence of breast cancer. This prompts further diagnostic tests or treatment.

ii. Negative result indicates that the CNN model did not find significant signs of breast cancer in the images. While reassuring, further screening may still be necessary in cases of high clinical suspicion.

4. RESULT

Our study showed that using a type of computer program called **2 Convolutional Neural Networks (CNNs)** is really good at finding signs **of breast cancer in** medical images. This CNN models for breast cancer detection benefit from advancements in architecture design, training methodologies, and interpretability techniques, leading to improved accuracy, reliability, and clinical relevance compared to traditional approaches. This new models learn from huge collections of images, teaching them what

cancer looks like. This helps them recognize cancer patterns better, even with smaller datasets. The program was able to correctly identify cancerous areas with high accuracy, usually more than 87% **5 of the time**, when we tested it on different sets of images. We also found that our program performed better than other methods that are currently used for breast cancer detection. The reason our program worked well is because **20 it can automatically learn** important **features from the** breast images without needing humans to point them out. This is thanks to the way CNNs are designed to understand images. Also, we had a lot of images for the program to learn from, which helped it become really good at its job. **23 In summary, our** study suggests that using CNNs **can be a** powerful tool for detecting breast cancer.

Fig. 4.1. Output

5. CONCLUSION

Breast cancer affects many women, with about 1 in 8 being diagnosed. Detecting it accurately can be tough **11 due to various** factors, leading to potential errors in diagnosis. Sometimes, tests can wrongly suggest cancer, leading to unnecessary procedures like biopsies. **19 To improve accuracy**, computer programs have been created to aid doctors in diagnosis. CNNs are adept at automatically learning hierarchical representations of data. This study focused on understanding how these computer programs work. They go through steps like preparing images, breaking them down, identifying important features, and using those features **4 to determine if cancer is present**. Once features are extracted, CNNs can classify **2 the likelihood of** these regions being indicative **of breast cancer. They** can distinguish between **benign and malignant** patterns with high accuracy aiding in early diagnosis and treatment planning. Texture is one such important feature that can indicate cancer cells. It's important to note that cancer isn't **21 just one disease but** a group **of many types**, making finding a universal cure impossible. Much research has been done, but a single drug to treat all cancers doesn't exist. CNNs offer a powerful tool for enhancing the accuracy, efficiency, and scalability **2 of breast cancer detection** from mammographic images, potentially leading to earlier detection and improved patient outcomes.

Future Scope

"The future potential ⁹ of Convolutional Neural Networks (CNNs) in detecting breast cancer appears promising. These networks act like highly intelligent detectives for images, capable of learning to identify subtle patterns in breast scans that may signify cancerous tissues. As technology progresses, CNNs can further enhance their ability to detect minuscule details or anomalous patterns that may elude human observation. Consequently, they hold the promise of aiding healthcare professionals ²⁴ in identifying breast cancer at earlier stages and with greater accuracy, potentially leading to life-saving interventions."

6. REFERENCES

- [1] ¹⁶ Zahra abdolali kazemi, "Diagnosis of breast cancer using image processing with SVM and KNN" Vol 13 no.1 May 2021 *International Journal of Engineering and Technology* 13(2):11
DOI:10.21817/ijet/2021/v13i2/211302004.

- [2] Saif Ali et.al "Identification of Cancer Using Image Processing Approaches" Vol. 9, ² *International Journal of Intelligent Information Systems*, July 2020
DOI: 10.11648/j.ijis.20200902.11.

- [3] Mutiullah et.al "Lung *Cancer Detection Using* Digital Image Processing Techniques: Vol. 38, ¹⁸ *Mehran University Research Journal Of Engineering & Technology* Journal, April 2021.

- [4] Yousif M. Y Abdallah et.al "Breast cancer detection using image enhancement and segmentation algorithms" *Biomedical Research* 2020 Volume 29 Issue 20.

- [5] Prannoy Giri et.al "Breast Cancer Detection using Image Processing Techniques" April 2020
An ² *International research journal of Computer Science and Technology*, DOI:
<http://dx.doi.org/10.13005/ojcs/10.02.19>

[6] Arpita Joshi and Dr. Ashish Mehta “Comparative Analysis of Various Machine Learning Techniques for Diagnosis of Breast Cancer” Vol. 10 Issue 07, July-2020 International Journal of Engineering Research & Technology (IJERT).

[7] David A. Omondiagbe, Shanmugam Veeramani and Amandeep S. Sidhu “Machine Learning Classification Techniques for Breast Cancer Diagnosis” (2020).

[8] Kalyani Wadkar, Prashant Pathak and Nikhil Wagh “Breast Cancer Detection Using ANN Network and Performance Analysis with SVM” International Journal of Computer Engineering and Technology, 10(3), pp. 75-86, (2020).

[9] Anji Reddy Vaka, Badal Soni and Sudheer Reddy “Breast Cancer Detection by Leveraging Machine Learning” ICT Express Volume 6, Issue 4, December 2020, Pages 320-324.

[10] Monika Tiwari, Rashi Bharuka, Praditi Shah and Reena Lokare “Breast Cancer Prediction using Deep Learning and Machine Learning Techniques” K. J. Somaiya Institute of Engineering and Information Technology (KJSIEIT) March 22, 2020.

Sources

1	https://www.ijert.org/breast-cancer-detection-using-machine-learning-techniques INTERNET 6%
2	https://www.hindawi.com/journals/jhe/2021/5528622/ INTERNET 3%
3	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7237065/ INTERNET 1%
4	https://nbcf.org.au/about-breast-cancer/detection-and-awareness/mammograms/ INTERNET 1%
5	https://my.clevelandclinic.org/health/diseases/17802-breast-calcifications INTERNET 1%
6	https://www.freecodecamp.org/news/form-validation-in-javascript/ INTERNET 1%
7	https://collegedunia.com/college/15260-shri-sant-gajanan-maharaj-college-of-engineering-ssgmce-shegaon INTERNET <1%
8	https://www.codecademy.com/article/what-are-ui-and-ux-design INTERNET <1%
9	https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/ INTERNET <1%
10	https://www.nature.com/articles/s41598-021-95519-0 INTERNET <1%
11	https://www.frontendmentor.io/articles/integrating-with-apis-a-beginners-guide-for-frontend-developers-u_xfEadBc- INTERNET <1%
12	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8607714/ INTERNET <1%
13	https://divyansh-singhal216.medium.com/machine-learning-project-using-svm-random-forest-and-knn-classifier-4c4981e3a589 INTERNET <1%
14	https://www.cancerresearchuk.org/443/about-cancer/breast-cancer/getting-diagnosed/finding-breast-cancer-early INTERNET <1%

15	https://pressroom.cancer.org/GlobalCancerStatistics2024 INTERNET <1%
16	https://www.researchgate.net/profile/Zahra-Abdolali-Kazemi-2/publication/351779950_Diagnosis_of_breast_cancer_using_image_processing_with_SVM_and_KNN/links/60a9776d299bf1031fc13dc8/Diagnosis-of-breast-cancer-using-image-processing-with-SVM-and-KNN.pdf INTERNET <1%
17	https://www.researchgate.net/publication/332446782_Lung_Cancer_Detection_Using_Digital_Image_Processing_Techniques_A_Review INTERNET <1%
18	https://www.muett.edu.pk/research/research-journals INTERNET <1%
19	https://neptune.ai/blog/building-ml-model-training-pipeline INTERNET <1%
20	https://www.datacamp.com/443/tutorial/tutorial-deep-learning-tutorial INTERNET <1%
21	https://alliancecancer.com/education-and-support/types-of-cancer/#:~:text=There are more than 100 different types of,of the skin is called basal cell carcinoma. INTERNET <1%
22	https://www.sitepoint.com/client-side-form-validation-html5/ INTERNET <1%
23	https://www.nature.com/articles/s41467-022-31535-6 INTERNET <1%
24	https://journals.lww.com/md-journal/fulltext/2024/01190/breast_cancer__a_review_of_risk_factors_and.67.aspx INTERNET <1%
25	https://www.sciencedirect.com/443/science/article/pii/S1746809421007382 INTERNET <1%

EXCLUDE CUSTOM MATCHES

OFF

EXCLUDE QUOTES

ON

EXCLUDE BIBLIOGRAPHY

ON