Proposal for a Research Project

Author : Munteanu Bianca-Ștefania

1. Title of the research project: Multimodal Diagnostic Tool for Breast Cancer Classification

2. Purpose of the Research Project

Breast cancer ranks as the most common cancer among women in developed countries. Its early detection, crucial for reducing death rates, heavily relies on medical imaging. Among various techniques, digital mammography is prominent but not infallible, as it misses around 20% of breast lesions. This shortfall necessitates additional biopsies, often on benign lesions. To enhance detection accuracy, other methods like ultrasound and magnetic resonance imaging are increasingly used alongside mammography. These complementary methods, using non-ionizing radiation, help refine diagnoses by integrating diverse imaging data, thereby improving the overall diagnostic accuracy in breast cancer cases [1]

Mammography, a standard tool for breast cancer screening, employs low-dose X-rays to capture breast images, helping identify potential signs of cancer such as masses and calcifications. However, its effectiveness diminishes in detecting tumors within dense breast tissue, a condition often linked with a higher cancer risk. This highlights the critical role of ultrasound as a supplementary modality. Breast ultrasound, low invasive and radiation-free, utilizes sound waves to detail the shape, border, and internal features of tumors, especially small, node-negative ones. Its lower cost and superior performance in dense breasts have made it increasingly vital in breast cancer diagnosis, notably for Asian women who commonly have denser breast tissue. When used alongside mammography, ultrasound significantly boosts cancer detection rates, emphasizing the necessity of a multimodal approach in breast cancer screening [2] [3].

The integration of mammogram and ultrasound data in breast cancer detection is a significant advancement in medical imaging. Mammography, primarily used for initial breast cancer (BC) screening, excels in identifying the presence of lesions. However, its limitation lies in cases where lesions are not palpable or visible in the mammogram. This is where ultrasound comes into play, with its strength in visualizing and localizing such elusive lesions. Notably, ultrasound is more adept at predicting the actual size of a tumor compared to clinical examinations and mammography alone. Therefore, by combining these two modalities, a more accurate and comprehensive assessment of the disease can be achieved, significantly improving the accuracy of both screening and diagnosis [4]

Building on the foundational understanding of the benefits and limitations of mammography and ultrasound in breast cancer detection, this project aims to develop a dual-modality, computer-aided system for more accurate breast tumor classification. The core purpose is to integrate and analyze data from both mammograms and ultrasounds using advanced artificial intelligence techniques. This system is envisioned to significantly enhance diagnostic accuracy by effectively distinguishing between benign and malignant tumors. By leveraging the complementary strengths of both imaging techniques, the project seeks to overcome the inherent limitations observed when these modalities are used independently. The goal is to provide a more nuanced and comprehensive analysis of breast tumors, thereby contributing to better-informed clinical decisions and improved patient outcomes. The development of this computer-aided system represents a step forward in personalized medicine, where technology and healthcare converge to offer tailored diagnostic solutions. It underscores the project's commitment to advancing breast cancer diagnostics through innovation, ultimately aiming to reduce the rate of unnecessary biopsies and to facilitate early and accurate cancer detection.

3. State of the art

While numerous studies have focused on single-modality classifiers, leveraging either mammography or ultrasound, there's a growing recognition of the limitations these single approaches present. This chapter will explore groundbreaking research that adopts a multimodal approach, integrating both mammography and ultrasound techniques. Such studies exemplify the shift towards using combined modalities, offering enhanced accuracy and sensitivity in cancer detection. The chapter will review key studies that demonstrate the potential of this integrated approach, significantly improving diagnostic capabilities compared to single-modality systems.

The research paper [2] presents an innovative method for breast cancer diagnosis by integrating mammography and ultrasound imaging techniques. Recognizing the limitations inherent in single-modality imaging for breast cancer detection, the authors propose a multimodal approach. This method aims to utilize the distinct diagnostic strengths of both mammography and ultrasound to enhance detection accuracy and sensitivity. Key to the study is the use of three distinct classifiers: K-Nearest Neighbors, Naive Bayes, and Support Vector Machine. These classifiers process texture and morphological features extracted from both mammography and ultrasound images. The novelty lies in the individual processing of these features through each classifier, followed by an ensemble method that integrates the outcomes using a majority voting system. This approach is designed to leverage the strengths of each classifier and each imaging modality, ultimately improving the accuracy of breast cancer detection. In terms of results, the study reports significant improvements in the diagnosis of breast cancer. The ensemble classification method achieved an accuracy rate of 88%. This research underscores the potential and importance of multimodal imaging in medical diagnostics. Breast cancer, being a complex disease, often requires insights from different imaging techniques. The method proposed in the study could have a significant impact on patient outcomes and treatment strategies, offering a more robust diagnostic tool. Furthermore, the authors advocate for continued research to refine this technique and explore its applicability in other medical imaging scenarios. They suggest that the future of multimodal imaging in medicine, particularly for diagnostic processes, is promising.

This study [4] presents a groundbreaking approach in breast cancer detection using a hybrid deep learning system. This system uniquely integrates a Convolutional Neural Network (CNN) with Long Short-Term Memory (LSTM) networks to analyze images from both mammograms and ultrasounds. The objective is to harness the strengths of both imaging modalities, combined with advanced machine learning techniques, to improve breast cancer detection. The research employed a novel dataset, which was augmented to include 1032 images, providing a robust base for the deep learning model. Enhancing the dataset through augmentation is crucial for training the model on a diverse range of scenarios. This step significantly boosts the model's ability to generalize, ensuring more accurate classifications of new and varied images. The hybrid model's performance was evaluated using several metrics, with a particular focus on classification accuracy and the AUC. Remarkably, the system achieved a classification accuracy of 99% and an AUC of 0.99. These results are significantly higher than those achieved by traditional single-modality systems, showcasing the potential of this dual-modality approach in clinical applications. Moreover, the study highlights the importance of combining different types of data and analytical techniques in medical imaging. The integration of CNN and LSTM networks exemplifies the synergy that can be achieved by combining spatial image analysis with the ability to recognize patterns over sequences of images.

The article [1] discusses a study aimed at enhancing breast cancer diagnosis through a computer-aided system that integrates textural data from ultrasound and digital mammography images. This system uses a pattern recognition system operating on graphics processing units, which allows for efficient parallel processing. The methodology involved analyzing breast lesions, verified histologically, in images marked by radiologists. The system's core is a probabilistic neural network classifier, whose performance was assessed through re-substitution and external cross-validation methods. Results indicated that combining ultrasound and mammography features improved the classification accuracy in differentiating malignant from benign lesions, compared to using individual modalities. The study concluded that this system improves the accuracy of breast lesion discrimination, can be updated on-site with new data, and serves as a valuable tool for clinical second opinions. The use of GPUs significantly accelerated processing, enhancing its practical applicability.

The article [5] presents an innovative approach to diagnosis by integrating mammography and ultrasound imaging. The core of the research revolves around the development of a novel Modality-Correlation Embedding Model (MCM). The authors aimed to address the limitations of traditional diagnostic methods, which often process mammography and ultrasound data separately, leading to potential inaccuracies. The MCM model is designed to deeply analyze and integrate the data from both imaging modalities. The innovation lies in the model's ability to recognize and utilize the inherent correlation between mammography and ultrasound data. By embedding these correlations into the diagnostic process, the model provides a more holistic and accurate assessment of breast tumors. To validate their approach, Xi et al. conducted extensive experiments using real-world clinical data. They compared the performance of their MCM model against traditional methods that treat mammography and ultrasound data independently. The results were significant. The MCM model demonstrated superior accuracy in tumor classification, outperforming existing methods. The authors thoroughly explored the practical implications of their findings, suggesting that the MCM model can be readily integrated into existing diagnostic workflows. Furthermore, the study delves into the technical aspects of the model, including its design and computational efficiency. The authors emphasized the use of advanced algorithms and computing techniques to ensure that the model is not only accurate but also efficient in processing large datasets.

4. Stages/Objectives of the Research Project

In this chapter, we outline the methodical approach of our research project. Each stage of the project is designed with specific objectives to ensure a comprehensive and innovative exploration of breast cancer classification. From the initial stages of data collection and preprocessing to the final validation and testing of our integrated models, our methodological framework is structured to address the complexities and challenges of combining mammogram and ultrasound data for enhanced diagnostic accuracy.

4.1. Data Collection and Collaboration

The initial stage involves the collection of a unique dataset comprising mammogram and ultrasound images. Given the absence of a public dataset with paired images, our objective is to collaborate with medical institutions or researchers who have access to such data. This stage is crucial for acquiring a diverse and representative dataset that forms the foundation of our research.

4.2. Preprocessing

This stage encompasses several key preprocessing steps, including resizing, normalization, contrast enhancement, noise reduction, artifact removal, data augmentation, and region of interest (ROI) extraction. The aim is to standardize and enhance the image quality, ensuring the dataset is primed for effective analysis by our AI models.

4.3. Model Development: Individual Modality Analysis

Here, we focus on developing and refining individual classification models for each imaging technique. This involves understanding the unique diagnostic capabilities and limitations of mammograms and ultrasounds separately, laying the groundwork for effective multimodal integration.

4.4. Data Fusion and Transformer Model Implementation

In this stage, we innovate a fusion technique to synergize the strengths of both imaging modalities. This includes the integration of multi-scale feature extraction, attention mechanisms, and transformer models, aiming to create a comprehensive system that enhances diagnostic accuracy.

4.5. Model Integration and Classification

The fused features are then processed through a sophisticated classification layer. This stage is pivotal in translating the rich, integrated data into a definitive diagnostic output, categorizing tumors as benign or malignant.

4.6. Validation and Performance Evaluation

The final stage is dedicated to rigorously testing and validating the integrated model. We aim to evaluate the model using key performance metrics like accuracy, sensitivity, specificity, AUC, PPV, and NPV. This stage is crucial for confirming the effectiveness and superiority of the multimodal approach over single-modality models.

Each stage of the project is carefully structured to address the complexities of combining mammogram and ultrasound data, with the overarching goal of significantly improving breast cancer classification methodologies. This comprehensive approach ensures that every aspect of the model development, from initial data collection to final validation, is meticulously planned and executed.

5. Research Methodology

This chapter details two potential data collection strategies: collaboration with local hospitals and requesting data from other researchers. It then delves into various preprocessing steps like resizing, normalization, contrast enhancement, noise reduction, and artifact removal, emphasizing their importance in preparing data for the convolutional neural network. The chapter also discusses data augmentation techniques and Region of Interest (ROI) extraction, highlighting their role in enhancing model performance. The latter sections focus on data fusion techniques for integrating mammogram and ultrasound data, including multi-scale feature extraction, an attention mechanism layer, feature fusion, the use of a transformer model, and the final classification layer. The chapter concludes with a detailed explanation of performance evaluation metrics, essential for assessing the effectiveness of the breast cancer classification system.

5.1 Dataset

The dataset is a fundamental component of our research methodology, presenting a unique challenge for our study. Currently, there are no publicly available datasets that contain paired mammogram and ultrasound images from the same patients. This presents two potential avenues for data acquisition:

- Collaboration with Local Hospitals
  We are considering partnering with local hospitals to collect mammogram and ultrasound images from the same patients. This collaboration would involve obtaining necessary ethical approvals and ensuring patient confidentiality and data

privacy. The process would entail careful coordination with medical professionals to acquire and anonymize the data, ensuring it is suitable for research purposes.

- Requesting Data from Other Researchers
  Another approach is to contact authors of relevant studies, such as those mentioned in the articles [4] and [1]. These studies have developed their own datasets from clinical settings. We plan to reach out to these researchers to request access to their datasets, offering potential for collaboration or citing their work in our research. For instance, the dataset described in the first article was developed from 31 patients at the All India Institute of Medical Sciences and comprises 86 images (43 mammograms and 43 ultrasound images). The second article details a dataset of digital mammography and ultrasound images from 62 patients, collected at the Delta Digital diagnostic center in Greece. Both datasets were collected with ethical approvals and could provide a valuable foundation for our study.

5.2 Preprocessing

Preprocessing is a critical step in ensuring that the input data is in a suitable form for the convolutional neural network (CNN) to process effectively. These steps are designed to enhance image quality, standardize input formats, and isolate key features, thereby laying the groundwork for accurate and efficient model performance. Each preprocessing technique, from resizing and normalization to contrast enhancement, noise reduction, and region of interest extraction, is chosen for its ability to refine the images and render them more conducive to successful machine learning analysis. The subsequent sections elaborate on each of these critical preprocessing steps, providing insights into their methodologies and justifications for their selection in our research approach.

5.2.1 Resizing

Resizing is the initial and fundamental preprocessing step in preparing our dataset for the convolutional neural network. Standardizing the dimensions to a uniform scale is essential for consistency and comparability. In our methodology, we will resize all images to a resolution of 512x512 pixels. This resolution is chosen to preserve sufficient detail that may be critical for accurate classification while ensuring that the images are suitably dimensioned for efficient processing by the CNN. Resizing not only facilitates the uniformity of the input layer but also optimizes computational efficiency. Care will be taken to maintain the aspect ratio of the images, preventing distortion that could potentially introduce artifacts or skew the representation of pertinent features. This step ensures that every image fed into the CNN starts from a standardized baseline, thus removing any bias that might arise from varying image sizes in the dataset.

5.2.2 Normalization

We will employ Min-Max normalization to standardize the pixel values of mammogram and ultrasound images, scaling them to a uniform range of [0, 1]. This linear preprocessing technique is selected for its ability to preserve the relationships among the original data values while adjusting the scale to a normalized range. By transforming the pixel intensities to a common scale, we ensure that each feature contributes equivalently to the model's input, thereby facilitating the CNN's learning process. The consistent input range is crucial for the effective training of the model, as it helps in accelerating the convergence and improving the stability of the gradient descent optimization throughout the learning phase [6].

5.2.3 Contrast Enhancement

Contrast enhancement, specifically through the use of Contrast-Limited Adaptive Histogram Equalization (CLAHE), is a vital preprocessing step in medical image analysis. CLAHE is designed to improve the visibility of important features within an image by adjusting the image contrast. The technique involves segmenting the image into smaller sections, known as tiles and subsequently performing histogram equalization on each of these segments. Unlike standard histogram equalization, CLAHE limits the contrast amplification to reduce the noise amplification typically seen with standard methods. As a result, it enhances the local contrast and brings out more detail without increasing noise levels, which is particularly beneficial for highlighting subtle nuances in mammograms and ultrasound images [7].

5.2.4 Noise Reduction

In our research, we will employ median filtering as a key noise reduction technique to enhance the quality of images. In this approach, the median value from a pixel's surrounding area is used to replace the pixel itself. The strength of median filtering lies in its ability to diminish noise while preserving the crucial edges within the images, a feature especially vital in medical imaging where clarity and detail are paramount. To evaluate the effectiveness of this technique, we will utilize established metrics such as root mean square error and standard deviation. These measures will ensure that the approach effectively reduces noise while also preserving the diagnostic integrity of the mammograms and ultrasound images [7].

5.2.5 Artifact Removal

This is an important step to ensure that only relevant image content is presented to the classification algorithms. Artifacts in medical images, such as labels, borders, or text annotations, can be a source of distraction and may lead to false interpretations by computer-aided diagnosis systems. These non-diagnostic elements are typically removed through a series of image processing techniques, including cropping, masking, or automated detection algorithms that can identify and exclude non-tissue regions. By focusing only on the diagnostically relevant areas, the effectiveness of subsequent image processing steps, such as feature extraction and classification, is significantly improved.

5.2.6 Data augmentation

In our research, data augmentation will play a crucial role in enhancing the robustness and generalizability of the model. To artificially expand our training dataset and introduce a wider variety of cases, we will implement several augmentation techniques on the mammogram and ultrasound images. Specifically, we will apply rotations at various angles (ranging from -25 to 25 degrees in 5-degree increments), zooming (up to 20% in and out), horizontal and vertical flipping, and translations (shifting the image by up to 10% of its original size in both the X and Y directions). These transformations will simulate different viewing conditions and patient positions, helping the model to learn from a more diverse set of image characteristics. By doing so, we aim to reduce overfitting and improve the model’s ability to perform accurately on unseen data, a crucial aspect for a reliable diagnostic tool. This methodical approach to data augmentation ensures that the model is exposed to a broad spectrum of variations, thereby enhancing its diagnostic performance in real-world scenarios.

5.2.7 Region of Interest (ROI) Extraction

In our study, the extraction of Regions of Interest (ROI) is a critical preprocessing step, designed to direct the focus of our model to the most relevant areas within the mammogram and ultrasound images. This process involves identifying and isolating specific sections of the images that are likely to contain key diagnostic information, such as potential tumor sites. By cropping these areas, we ensure that the model's analysis is concentrated on the parts of the image that are most informative for breast cancer classification. This targeted approach not only enhances the efficiency of the model by reducing the computational load but also improves the accuracy of the analysis by minimizing the influence of irrelevant background information. The precision in ROI extraction is paramount, as it directly impacts the quality of the features that will be fed into the convolutional neural network for further processing. Utilizing advanced image processing techniques, the ROIs will be accurately delineated, ensuring that the subsequent stages of the model have access to the most critical visual information necessary for reliable tumor classification.

5.3 Data fusion techniques for integrating mammogram and ultrasound data

5.3.1. Input Layer

The data fusion process begins at the input layer, where pairs of mammogram and ultrasound images are fed into the system. This first step is important as it sets the stage for the analysis. The input layer is designed to accommodate the distinct characteristics of each modality – mammograms providing detailed structural information and ultrasounds offering complementary textural and morphological insights. Care is taken to ensure that the input images are aligned and correspond to the same anatomical regions, facilitating effective subsequent fusion. By starting with pairs of images, the model is provided with a comprehensive view from the outset, enhancing its ability to detect and classify breast tumors accurately.

5.3.2. Multi-scale Feature Extraction

In this phase of our data fusion model, we utilize a specially designed convolutional neural network (CNN) architecture to process mammogram and ultrasound images at multiple scales. This involves a variety of layers including convolutional layers with different kernel sizes to capture a wide range of features, both fine and broad. Pooling layers are used for dimensionality reduction, enhancing focus on essential features while Inception modules allow for parallel processing at various scales. Dilated convolutional layers are included to broaden the receptive field, capturing more contextual information, while depthwise separable convolutional layers efficiently capture important features with lower computational overhead. To ensure consistent feature integrity throughout the network, we integrate skip connections, which help in preserving important information across layers. Batch normalization layers are strategically placed after convolutional layers to stabilize and accelerate the learning process, enhancing the network's performance. This layer composition in our CNN architecture is pivotal to achieving comprehensive and detailed feature extraction from medical images, a fundamental requirement for the accurate classification of tumors. This multi-scale feature extraction process is designed not just for precision, but also for efficiency, ensuring the system is both accurate and scalable for practical medical applications.

5.3.3. Attention Mechanism Layer

After the multi-scale feature extraction, our model incorporates an attention mechanism layer. This layer plays a pivotal role in enhancing the model's focus on the most significant regions of the mammogram and ultrasound images. By employing an attention mechanism, the network learns to assign varying degrees of importance to different areas of the image, effectively highlighting regions that are more likely to contain diagnostically relevant information. This selective focus aligns with how a radiologist might prioritize certain features or anomalies within an image more closely than others. The attention layer dynamically adjusts weights across the image, ensuring that the model's subsequent processing and analysis give priority to these key areas. This focus on critical regions not only improves the accuracy of tumor detection and classification but also makes the model's decision-making process more interpretable and aligned with clinical assessment practices.

5.3.4. Feature Fusion

Following the attention mechanism, the next crucial step in our model is the fusion of features extracted from both types of images. This feature fusion stage is key to integrating the unique insights provided by each imaging modality. We employ sophisticated fusion techniques that go beyond simple concatenation, aiming to merge the features in a way that captures both the complementary and the distinctive characteristics of mammogram and ultrasound data. This could involve advanced strategies like feature concatenation followed by dimensionality reduction, or more complex algorithms that intelligently combine features to enhance their representational power. The goal of this fusion process is to create a comprehensive feature vector that encapsulates a complete picture of the underlying tissue characteristics. By doing so, we leverage the strengths of both imaging types, ensuring that the final feature set used for classification is robust, detailed, and highly informative. This integrative approach is expected to significantly improve the model's ability to accurately classify breast tumors, benefiting from the combined diagnostic value of mammogram and ultrasound features.

5.3.5. Transformer Model

Once the features from mammogram and ultrasound images are fused, the next step in our model involves passing these integrated features through a transformer model. The transformer, renowned for its effectiveness in processing sequential data, uses self-attention mechanisms to analyze the feature sequence. This allows the model to weigh the importance of each feature within the context of the entire sequence, enhancing its ability to discern patterns and relationships that might be critical for accurate classification. The self-attention mechanism in the transformer model provides a dynamic way of emphasizing certain features over others based on their relevance to the task at hand, which in our case is the classification of breast tumors. This approach enables the model to focus on the most informative aspects of the combined feature set, potentially uncovering subtle but crucial diagnostic signals that might be missed by traditional CNN architectures.

5.3.6. Classification Layer

The final layer of our model's processing pipeline is the classification layer, where the refined features, processed through the transformer model, are ultimately used for the critical task of tumor classification. In this layer, the combined and attention-enhanced features from mammogram and ultrasound images are inputted to make the final determination: classifying the tumor as benign or malignant. This is achieved through a series of fully connected layers and activation functions that interpret the feature set's intricate patterns, translating them into a definitive diagnostic decision. This layer's architecture is crucial, as it directly impacts the model's accuracy and reliability.

5.3.7. Output

The final stage in our model's workflow is the output, where the system conveys its conclusive analysis. This output is the definitive classification of the breast tumor, determined as either benign or malignant.

5.4 Performance evaluation metrics

In our research, we will employ the same performance evaluation metrics as the article [2] to assess the efficacy of our breast cancer classification system, which are essential for determining the system's effectiveness.

Note the following notations:
TP stands for instances correctly identified as positive.
TN denotes instances accurately recognized as negative.
FP represents cases mistakenly classified as positive.
FN refers to cases incorrectly labeled as negative.

1. Accuracy = $\frac{TP+TN}{TP+TN+FP+F} * 100\%$

This represents the proportion of correctly identified cases in the total sample. It's the sum of true positive and true negative predictions, divided by the total number of predictions, then multiplied by 100.

2. Sensitivity = $\frac{TP}{TP+FN} * 100\%$

Also known as recall, this metric quantifies the model's accuracy in identifying malignant cases. It's the ratio of true positive predictions to all actual positive cases.

3. Specificity = $\frac{TN}{TN+FP} * 100\%$

This measure reflects the accuracy in identifying benign cases, calculated by the proportion of true negative predictions among all actual negative cases.

4. Area Under the Curve (AUC)

AUC measures the model's ability to distinguish between classes, represented by the area under the ROC curve plotting sensitivity against 1-specificity.

5. Positive Predictive Value, PPV = $\frac{TP}{TP+FP} * 100\%$

This indicates the proportion of positive test results that are true positives, showing the likelihood that a positive diagnosis is correct.

6. Negative Predictive Value, NPV = $\frac{TN}{TN+F} * 100\%$

NPV is the proportion of negative test results that are true negatives, indicating the likelihood that a negative diagnosis is accurate.

6. Anticipated Results

In this research, we anticipate a range of significant outcomes that will not only validate the effectiveness of our proposed model but also contribute to the broader field of medical diagnostics. We expect the following specific performance metrics:

- Enhanced accuracy in tumor classification: We anticipate achieving an accuracy rate significantly higher than current standards, potentially exceeding 90%. This improvement in accuracy is expected due to the synergistic effect of combining mammogram and ultrasound imaging with advanced AI techniques like multi-scale feature extraction and transformer models.

- Improved reliability in difficult to diagnose cases: Our model aims to demonstrate improved reliability, especially in challenging diagnostic scenarios. We expect sensitivity (true positive rate) to be above 85%, ensuring that the model effectively identifies malignant tumors, and specificity (true negative rate) to also be above 85%, confirming its ability to correctly dismiss benign cases.

- Potential for reducing false positives/negatives: One of the model's key objectives is to reduce the incidence of false positives and negatives. We anticipate a high Positive Predictive Value (PPV) and Negative Predictive Value (NPV), both potentially exceeding 80%, which would significantly minimize the rate of misdiagnoses.

- Area Under the Curve (AUC) expectations: For the AUC, which measures the model's ability to differentiate between benign and malignant tumors, we aim for a value close to 0.9 or above, indicating excellent diagnostic ability.

- Broader implications for multimodal diagnostic tools in medicine: Beyond breast cancer diagnosis, the success of this model could set a precedent for multimodal diagnostic tools in various medical fields. The implications of achieving these high-performance metrics could pave the way for AI-driven diagnostics to become a standard in personalized medicine, enhancing the accuracy and reliability of disease detection and treatment plans across a multitude of medical conditions.

These anticipated results, characterized by high accuracy, sensitivity, specificity, AUC, PPV, and NPV, will not only validate our model's effectiveness but also demonstrate the profound impact of integrating multimodal data and advanced AI techniques in medical diagnostics.

7. Proposed Team and Budget Estimation

Team Composition

1. Project Lead (1) - Oversees the entire project, ensures adherence to objectives, and communicates with stakeholders.
2. Data Scientists (2) - Specialize in AI model development, data analysis, and algorithm optimization.
3. Medical Imaging Specialists (2) - Provide expertise in mammography and ultrasound imaging, and assist in data interpretation.
4. Software Developers (2) - Develop and maintain the software infrastructure for the AI models.
5. Research Assistants (2) - Aid in data collection, preprocessing, and administrative tasks.
6. Ethics and Compliance Officer (1) - Ensures all research adheres to ethical standards and data privacy laws.

Budget Estimation:

- Personnel Costs: 250.000€ (salaries for team members).
- Data Acquisition and Processing: 30.000€ (costs for dataset collection, storage, and preprocessing tools).
- Software and Hardware: 50.000€ (computational resources, software licenses).
- Contingency Fund: 20.000€.

Total Estimated Budget: 350.000€.

References

[1] K. P. Sidiropoulos, S. A. Kostopoulos, D. T. Glotsos, E. I. Athanasiadis, N. D. Dimitropoulos, J. T. Stonham și D. A. Cavouras, „Multimodality GPU-based computer-assisted diagnosis of breast cancer using ultrasound and digital mammography images,” *International journal of computer assisted radiology and surgery,* vol. 8, pp. 547--560, 2013.

[2] P.-H. Tsui, J. Cong, B. Wei, Y. He, Y. Yin și Y. Zheng, „A selective ensemble classification method combining mammography images with ultrasound images for breast cancer diagnosis,” *Computational and mathematical methods in medicine,* vol. 2017, 2017.

[3] A. Sahu, P. K. Das și S. Meher, „High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets,” *Biomedical Signal Processing and Control,* p. 104292, 2023.

[4] K. Atrey, B. K. Singh, N. K. Bodhey, și R. Bilas Pachori, „Mammography and ultrasound based dual modality classification of breast cancer using a hybrid deep learning approach,” *Biomedical Signal Processing and Control,* vol. 86, p. 104919, 2023.

[5] X. Xi, W. Li, B. Li, D. Li, C. Tian și G. Zhang, „Modality-correlation embedding model for breast tumor diagnosis with mammography and ultrasound images.,” *Computers in biology and medicine,* vol. 150, 2022.

[6] B. K. Singh, K. Verma și A. Thoke, „Investigations on impact of feature normalization techniques on classifier's performance in breast tumor classification,” *International Journal of Computer Applications,* vol. 116, nr. 19, 2015.

[7] H. Avcı și J. Karakaya, „A Novel Medical Image Enhancement Algorithm for Breast Cancer Detection on Mammography Images Using Machine Learning,” *Diagnostics,* vol. 13, nr. 3, 2023.