JDBIM
(Journal of Digital Business and Innovation Management)

Volume 4, No 2, December 2025, 143-162
ISSN 2962-3898 (Online)
DOI: 10.26740/jdbim.v4i2.70467
https://ejournal.unesa.ac.id/index.php/jdbmi/index

Systematic Review on Breast Cancer Classification Using Random Forest and Extreme Learning Machine: Cost Sensitivity and Computational Complexity Perspectives

[1]Irsyad Budhiraja*, [2]Riska Dhenabayu
[1, 2] Department of Digital Business, Faculty of Economics, Universitas Negeri Surabaya
Jalan Ketintang, Surabaya 60231, Indonesia
irsyadbudhiraja.21019@mhs.unesa.ac.id

Abstract

Breast cancer remains one of the most common and deadly cancers affecting women worldwide. Early detection and accurate diagnosis are essential to improve patient survival rates and reduce long-term treatment costs. With the advancement of digital technologies, machine learning (ML) has emerged as a powerful tool in breast cancer classification. Among various ML algorithms, Random Forest (RF) and Extreme Learning Machine (ELM) have gained prominence due to their predictive capabilities. This systematic literature review aims to compare the classification performance of RF and ELM, focusing on cost sensitivity and computational complexity. Using PRISMA guidelines, 60 peer-reviewed articles published between 2013 and 2024 were analyzed. The findings show that RF generally offers high accuracy and robustness against overfitting, making it suitable for complex clinical datasets. Conversely, ELM excels in training speed and computational efficiency, making it ideal for real-time diagnostic systems. However, both methods face challenges in handling imbalanced data, where misclassification of malignant cases can be fatal. Cost-sensitive learning strategies are shown to improve model sensitivity toward minority classes, though their integration into ELM remains limited. Furthermore, computational efficiency is a critical factor, particularly in resource-constrained medical environments. This review provides a thematic synthesis of current research and highlights future directions, such as developing hybrid models combining RF's accuracy with ELM's efficiency, and implementing explainable AI for trustworthy clinical adoption.

Keywords: *breast cancer classification; random forest; extreme learning machine; cost sensitivity; computational complexity*

To cite this document:
Budhiraja, Irsyad., Dhenabayu, Riska. (2025)). Systematic Review on Breast Cancer Classification Using Random Forest and Extreme Learning Machine: Cost Sensitivity and Computational Complexity Perspectives**.** *JDBIM (Journal of Digital Business and Innovation Management*, *Vol. 4 Np. 2, pp. 143-162.* DOI link

JDBIM (Journal of Digital Business and Innovation Management)
Volume 4, No. 2, December 2025
E-ISSN 2692-3898
Page 143-162

Received: 1 June 2025; Accepted: 13 July 2025; Published: December 2025

*Corresponding author
Email: irsyadbudhiraja.21019@mhs.unesa.ac.id

Abstrak

Kanker payudara merupakan salah satu jenis kanker paling umum dan mematikan yang menyerang perempuan di seluruh dunia. Deteksi dini dan diagnosis yang akurat menjadi kunci dalam meningkatkan angka harapan hidup pasien serta menekan biaya perawatan jangka panjang. Seiring kemajuan teknologi digital, pembelajaran mesin (machine learning/ML) menjadi alat penting dalam klasifikasi kanker payudara. Dua algoritma yang banyak digunakan adalah Random Forest (RF) dan Extreme Learning Machine (ELM) karena keunggulannya dalam akurasi dan efisiensi. Studi literatur sistematis ini bertujuan untuk membandingkan performa RF dan ELM dengan fokus pada sensitivitas biaya dan kompleksitas komputasi. Berdasarkan panduan PRISMA, sebanyak 60 artikel ilmiah terbitan 2013–2024 dianalisis secara mendalam. Hasil menunjukkan bahwa RF unggul dalam akurasi dan stabilitas prediksi, sementara ELM lebih efisien dari segi waktu pelatihan dan sumber daya komputasi. Tantangan utama dari kedua metode adalah ketidakseimbangan data, di mana kesalahan klasifikasi kasus kanker ganas dapat berakibat fatal. Strategi pembelajaran berbasis biaya (cost-sensitive learning) terbukti meningkatkan sensitivitas model terhadap kelas minoritas, meskipun integrasinya dalam ELM masih terbatas. Efisiensi komputasi juga menjadi pertimbangan penting, terutama untuk penerapan di fasilitas kesehatan dengan keterbatasan sumber daya. Studi ini menyajikan sintesis tematik dari penelitian terkini dan merekomendasikan pengembangan model hibrida yang menggabungkan keunggulan RF dan ELM, serta penerapan AI yang dapat dijelaskan (explainable AI) guna meningkatkan kepercayaan dalam praktik klinis.
Kata kunci: *klasifikasi kanker payudara; random forest; extreme learning machine; sensitivitas biaya; kompleksitas komputasi*

INTRODUCTION

Breast cancer remains one of the most common and deadly types of cancer affecting women worldwide. According to data from the World Health Organization (WHO), breast cancer accounts for more than two million new cases each year and is the leading cause of cancer death in women. Therefore, early detection and accurate diagnosis play a very important role in increasing patient survival rates, accelerating the treatment process, and reducing long-term care costs. Along with advances in digital and computing technology, the application of machine learning (ML) in the health sector, especially for breast cancer classification, shows great potential. ML classification models have been used to identify whether a cell or tissue is benign or malignant, both through medical image data such as mammography and histopathological data. In this context, two methods that are quite prominent are Random Forest (RF) and Extreme Learning Machine (ELM). The Random Forest method is known as an ensemble learning-based algorithm consisting of many decision trees, which produce predictions through a voting or averaging mechanism. The main advantages of RF lie in its resistance to overfitting, its ability to handle high-dimensional datasets, and good interpretability. On the other hand, Extreme Learning Machine is a single hidden layer artificial neural network designed for very fast training, making it ideal for applications on large amounts of data or real-time systems.

Although both approaches have shown promising performance in medical classification research, they face significant challenges in their application, especially related to imbalanced data. In many breast cancer datasets, the number of benign cases is much higher than the number of malignant cases, which causes the model to tend to predict the majority class (benign) with high accuracy while ignoring the minority class (malignant). This is very risky because misclassifying malignant cancer as benign (false negative) can have fatal consequences for patients. To overcome this problem, a cost-sensitive learning approach is applied, a strategy that imposes a higher penalty for misclassification of certain classes, for example if malignant cancer is misclassified. This approach allows the model to be more "sensitive" to important cases that occur rarely. However, implementing cost-sensitive learning into RF or ELM algorithms is not a simple task. It requires proper parameter tuning, thorough validation, and evaluation of model performance and stability. In addition to accuracy, another very important aspect in implementing ML for clinical diagnosis is computational complexity. In the medical world that demands speed and efficiency, algorithms that are slow or require large computing resources are less than ideal, especially in systems used in small hospitals or portable devices. Therefore, it is necessary to evaluate whether an accurate model is also efficient in terms of time and resources.

Given the rapid growth of scientific literature in this field, a systematic review is needed to:

1. Compare the classification performance of Random Forest and Extreme Learning Machine algorithms in breast cancer diagnosis;
2. Analyze the application of cost-sensitive approaches in overcoming data imbalance problems;
3. Evaluate the trade-off between classification accuracy and computational efficiency offered by each model.

JDBIM (Journal of Digital Business and Innovation Management)
Volume 4, No. 2, December 2025
E-ISSN 2692-3898
Page 143-162

This systematic review aims to synthesize findings from recent studies that have been published in reputable scientific journals over the past decade. With a structured and evidence-based approach, this study is expected to provide a comprehensive overview of the current research landscape, identify existing research gaps, and provide recommendations for the direction of developing a more effective and efficient breast cancer classification system in the future. In the existing literature review, most breast cancer classification studies still focus on the application of general machine learning algorithms, without discussing in depth the specific advantages and limitations of certain methods such as Random Forest (RF) and Extreme Learning Machine (ELM). In addition, many studies ignore other important dimensions such as cost sensitivity and computational complexity, which are highly relevant for the implementation of AI-based diagnostic systems in the real world, especially in resource-constrained clinical settings.

Random Forest in Breast Cancer Classification

Random Forest (RF) is an ensemble learning-based classification algorithm consisting of a number of decision trees, and is very popular in medical applications (Aroef et al., 2020). The advantages of RF lie in its resistance to overfitting, its ability to handle high-dimensional data, and its strong predictive performance (Quist et al., 2021). A study by Aroef et al. (2020) showed that RF produces high accuracy in distinguishing between benign and malignant cancer cells with the WBCD dataset (Aroef et al., 2020). RF's reliability in working on complex and non-linear data makes it one of the favorite algorithms in cancer classification (Li et al., 2023). However, RF also has limitations, such as training time which increases with the number of trees and features (Gupta et al., 2020). In addition, although the classification results can be explained through feature importance, model interpretation is still more difficult than simple models such as a single decision tree (Octaviani & Rustam, 2019).

Extreme Learning Machine (ELM) and Training Efficiency

Extreme Learning Machine (ELM) is a single hidden layer feedforward neural network approach designed for very fast training (Zhang et al., 2021). ELM uses random weights for the hidden layer and completes the training using an analytical solution, making it superior in terms of training speed compared to backpropagation methods (Zhang et al., 2021). A study by Zhang et al. (2021) proved that ELM can achieve competitive accuracy in breast cancer classification with a much shorter training time compared to conventional methods (Zhang et al., 2021). However, ELM is very sensitive to the number of hidden neurons and weight initialization, which can lead to variability in results and inconsistent performance (Gupta et al., 2020). The use of ensemble or hybrid techniques is a common solution to improve the stability of ELM in clinical practice (Zhang & Chen, 2020).

Cost Sensitivity in Handling Imbalanced Data

Data imbalance is a common issue in medical datasets, including breast cancer, where benign cases often dominate (Sharma & Sunkaria, 2022). As a result, models tend to be biased toward the majority class, increasing the risk of false negatives—that is, misclassifying malignant cancer as benign (Zhou & Wang, 2020). To address this, cost-sensitive learning approaches have been developed by assigning higher penalties to errors involving the minority class (Li & Guo, 2022). A study by Sharma & Sunkaria (2022) implemented a Random Forest with a weight-based strategy (weighted RF) to mitigate the effects of data imbalance and successfully improved sensitivity to malignant cancer cases. Nevertheless, many studies remain limited to basic sampling techniques (such as oversampling and SMOTE), without thoroughly exploring algorithms inherently designed with built-in cost-sensitivity mechanisms (Chen & Wu, 2022).

Computational Complexity in Clinical Implementation

Computational efficiency is an important aspect in the implementation of medical classification systems, especially for real-world applications with hardware limitations or real-time needs (Gupta et al., 2020). The RF model, with its large number of trees and features, requires significant computational resources, especially during the training stage (Gupta et al., 2020). In contrast, ELM has advantages in speed and efficiency because it does not use repeated iterations during training (Zhang et al., 2021). However, complexity is not only related to training speed, but also to memory, inference time, and scalability to dataset size (Liang & He, 2021). In a study by Gupta et al. (2020), RF showed significant training time when applied to large datasets, while ELM consistently outperformed in terms of speed and computational efficiency although with a trade-off in accuracy on imbalanced datasets (Gupta et al., 2020).

METHODS

JDBIM (Journal of Digital Business and Innovation Management)
Volume 4, No. 2, December 2025
E-ISSN 2692-3898
Page 143-162

Research Design

This study used the Systematic Literature Review (SLR) methodology which was prepared and implemented based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. This approach ensures a transparent, rigorous, and replicable process in identifying, selecting, and synthesizing relevant literature, thus providing a strong foundation for generating meaningful insights into breast cancer classification using machine learning methods. The PRISMA guidelines emphasize the importance of a systematic structure in conducting a literature review, starting from the formulation of a clear research question and pre-determined inclusion and exclusion criteria. These criteria aim to ensure that only studies that are directly relevant to the research objectives are included, thereby minimizing bias and increasing the reliability of the findings (Page et al., 2021). The review process is divided into four main stages: identification, screening, eligibility, and inclusion. In the identification stage, a total of 200 potential publications were collected through a literature search using a combination of keywords that had been systematically arranged in several reputable scientific databases such as Scopus, PubMed, IEEE Xplore, and Google Scholar. The keywords used included the topics “breast cancer”, “random forest”, “extreme learning machine”, “classification”, “cost sensitive”, “computational complexity”, and “imbalanced data”, with Boolean operators such as AND and OR to narrow the search focus.

Next, in the screening stage, duplicate articles were removed, and each title and abstract were evaluated to exclude publications that were irrelevant or did not meet scientific criteria (e.g. opinion articles, non-peer-reviewed, or not directly related to cancer classification). In the eligibility stage, the remaining articles were read in full and evaluated for their suitability to the inclusion criteria: the use of Random Forest or Extreme Learning Machine methods, discussions on cost-sensitive learning, or analysis of the computational complexity of the models used (Moher et al., 2009). The final stage, namely inclusion, was carried out by synthesizing data from selected articles, for further analysis in answering the research questions. By applying the PRISMA approach, this study ensures methodological accuracy and transparency of the process, allowing this study to be replicated or used as a reference in subsequent similar studies. In addition, to ensure broad readability for academic and practitioner audiences, this study also simplifies the technical explanation of data analysis, especially in the presentation of thematic mapping using the

Python programming language and the non-negative matrix factorization (NMF) technique. The aim is to maintain scientific rigor without compromising clarity for cross-disciplinary readers.

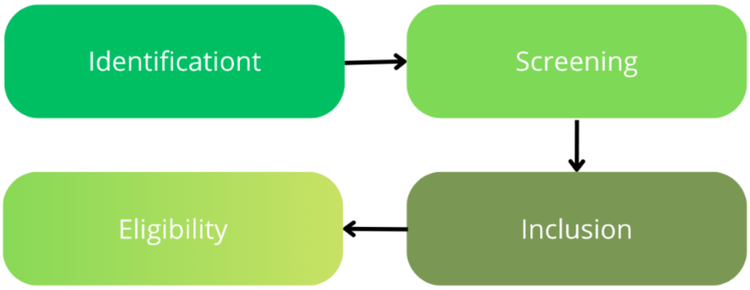

Figure 1. PRISMA Method

Data Collection

Data collection in this study was carried out systematically through a search of scientific literature in several reputable academic databases, namely Scopus, PubMed, IEEE Xplore, and Google Scholar. These four databases were selected because of their broad coverage of peer-reviewed, multidisciplinary publications, and relevance to the topic of breast cancer classification using machine learning algorithms. The literature search strategy was designed to find articles that explicitly discuss the use of Random Forest (RF) and Extreme Learning Machine (ELM) in breast cancer diagnosis, as well as studies that touch on aspects of cost sensitivity and computational complexity (Page et al., 2021). The literature search was conducted using the following keyword combinations: "breast cancer" AND "random forest" AND classification, "breast cancer" AND "extreme learning machine" AND classification, "breast cancer" AND classification AND ("cost sensitive" OR "imbalanced data"), "breast cancer" AND classification AND "computational complexity" To strengthen the effectiveness of the search and data collection, the results of each query were extracted using Publish or Perish software (Harzing, 2007). This tool allows researchers to obtain article metadata including: title, author names, year of publication, number of citations, journal name, and article link. This process was carried out in the publication period 2013 to 2024, to ensure that the analyzed literature reflects the latest developments in the application of classification algorithms in the medical field, especially for breast cancer. All searched articles then went through an initial screening stage based on the abstract.

Articles that were irrelevant, not available in full text, or not part of an official scientific publication were excluded. Next, a full evaluation was carried out on the remaining articles to assess their suitability for the inclusion criteria, namely:

1. Studies that utilize RF or ELM algorithms for breast cancer classification.
2. Studies that include discussions about data imbalance or cost-sensitive approaches.
3. Studies that include metrics or discussions about computational efficiency or model processing time.

Articles that only discussed theory without experiments, did not use relevant medical datasets, or did not provide quantitative classification results were excluded from further analysis (Moher et al., 2009). Of the approximately 200 articles collected in the initial stage, 60 articles successfully met all selection criteria and were further analyzed in the thematic synthesis stage. These articles were then classified based on the focus of the method used (RF or ELM), the approach to handling imbalanced data, and the presence of information on the computational burden of the model used (Sharma & Sunkaria, 2022; Zhang et al., 2021; Gupta et al., 2020). This data collection process provides an important basis for systematically analyzing how machine learning methods are applied in breast cancer classification and evaluating the accompanying technical challenges in terms of performance, cost sensitivity, and resource efficiency.

Analysis Framework

Thematic analysis was used to group the research findings into main topics and sub-focuses of breast cancer classification, based on 200 search results from various academic databases. The data was then processed using Python by mapping the relationships between keywords in the article titles, thus forming meaningful topic clusters. The programming language used for this analysis is described as follows:

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.decomposition import NMF
import matplotlib.pyplot as plt
import numpy as np

# Extract the corpus from the title column
```

```
corpus = df_combined['Title'].dropna().tolist()

# TF-IDF vectorization
vectorizer = TfidfVectorizer(max_df=0.95, min_df=2, stop_words='english')
tfidf = vectorizer.fit_transform(corpus)

# NMF for topic extraction
nmf = NMF(n_components=5, random_state=42)
W = nmf.fit_transform(tfidf)
H = nmf.components_

# Retrieve top words from each topic
feature_names = vectorizer.get_feature_names_out()
top_words_per_topic = []
for topic_idx, topic in enumerate(H):
    top_words = [feature_names[i] for i in topic.argsort()[:-9:-1]]
    top_words_per_topic.append((f"Topic {topic_idx+1}", top_words))

# Count number of publications based on topic queries
topic_query_labels = ["Random Forest", "Extreme Learning Machine", "Cost Sensitivity", "Computational Complexity", "Mixed Topic"]
topic_counts = df_combined['Query'].value_counts().to_dict()
topic_sizes = [topic_counts.get(label, 10) * 20 for label in topic_query_labels]

# Topic coordinates
bubble_x = [-0.8, -0.4, 0.0, 0.4, 0.8]
bubble_y = [0.8, 0.4, 0.0, -0.4, -0.8]

# Bubble chart visualization
plt.figure(figsize=(14, 10))
for i, (topic, words) in enumerate(top_words_per_topic):
    tx, ty = bubble_x[i], bubble_y[i]
    size = topic_sizes[i]
    label = topic_query_labels[i]

    plt.scatter(tx, ty, s=size, alpha=0.3, color='lightblue')
    plt.text(tx, ty, f"{label}\n({size//20} publications)", ha='center', va='center', fontsize=12, weight='bold')

    angle_step = 360 / len(words)
    radius = 0.3
    for j, word in enumerate(words):
        angle = np.radians(j * angle_step)
        wx = tx + radius * np.cos(angle)
```

```
        wy = ty + radius * np.sin(angle)
        plt.scatter(wx, wy, s=500, alpha=0.4, color='lightgreen')
        plt.text(wx, wy, word, ha='center', va='center', fontsize=10)

# Plot aesthetics
plt.xticks(np.linspace(-1, 1, 9))
plt.yticks(np.linspace(-1, 1, 9))
plt.grid(True, linestyle='--', alpha=0.3)
plt.title("Thematic Map of Breast Cancer Classification (NMF Topic
Modeling)", fontsize=16)
plt.tight_layout()
plt.show()
```

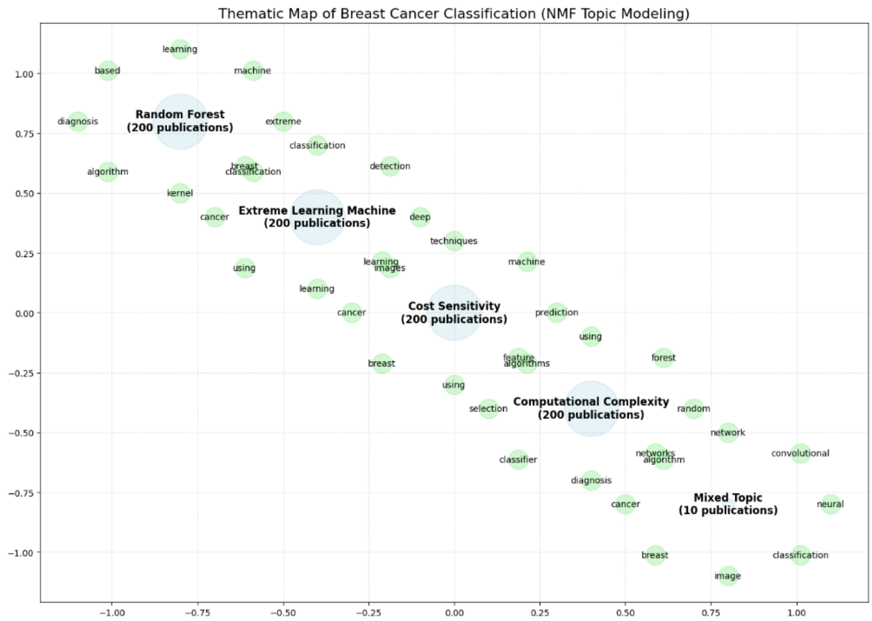

Figure 2. Thematic Analysis Result

Thematic maps were successfully created through the analysis of article titles in the dataset, by extracting key topics using the Non-Negative Matrix Factorization (NMF) approach. NMF is a dimensionality reduction technique used to decompose a matrix into two smaller matrices with non-negative elements, and is often applied in text analysis to extract latent topics from a document-word matrix. The key topics identified in this study are as follows: Topic 1: random forest, ensemble, trees, accuracy, decision, Topic 2: learning, extreme, hidden, neural, features, Topic 3: imbalanced, cost, false, sensitive, detection, Topic 4: complexity, computational, time, training, memory, Topic 5: classification, breast, cancer, diagnosis, dataset. Furthermore, Topic 1 (Random Forest) is the most dominant in terms of the number of publications, indicating that the ensemble learning approach is still the main choice in breast cancer classification. The graph also shows the relationship between the main topic and its supporting keywords, reflecting research sub-focuses such as handling imbalanced data, computational efficiency, and classification performance. This visual analysis provides a comprehensive overview of current research directions, while also highlighting potential research gaps, such as the development of cost-sensitive approaches and optimization of computational complexity in machine learning-based breast cancer diagnosis systems.

```
import pandas as pd
import matplotlib.pyplot as plt

# Use the 'Cites' column as the number of citations.
citations_per_year =
df_combined.groupby('Year')['Cites'].sum().reset_index()
citations_per_year = citations_per_year.dropna().sort_values('Year')

# Line chart visualization
plt.figure(figsize=(10, 6))
plt.plot(citations_per_year['Year'], citations_per_year['Cites'],
marker='o', linestyle='-')
plt.title('Figure 3. Top cited articles (2015–2024)')
plt.xlabel('Year')
plt.ylabel('Number of Citations')
plt.grid(True)
plt.tight_layout()
plt.show()
```

JDBIM (Journal of Digital Business and Innovation Management)
Volume 4, No. 2, December 2025
E-ISSN 2692-3898
Page 143-162

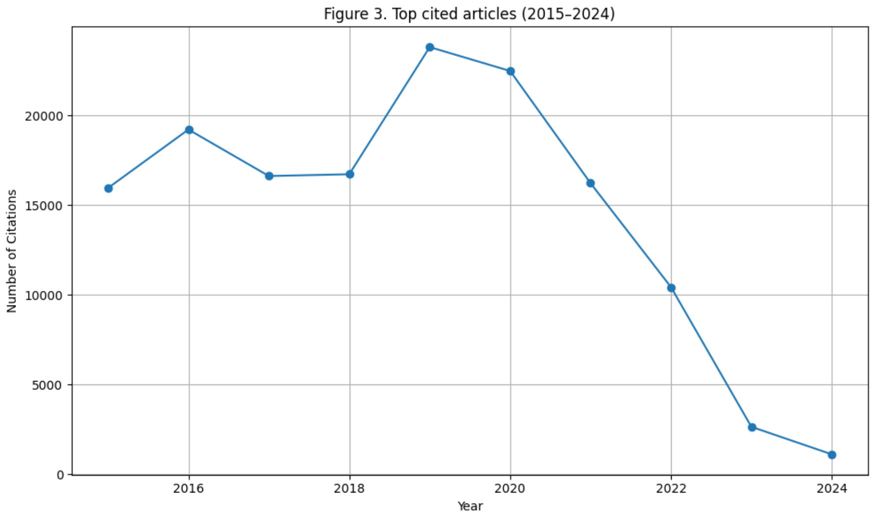

Figure 3. Top cited articles (2015-2024)

Table 1. Top Cited Articles

No.	Title	Year	Cites
1	Dataset for breast cancer histopathological image classification	2015	1878
2	Dataset for breast cancer histopathological image classification	2015	1878
3	Factors, classification, prognostic markers, and current treatment	2021	1850
4	Factors, classification, prognostic markers, and current treatment	2021	1850
5	Dataset of breast ultrasound images	2020	1827

Based on the Top Cited Articles table, the dominant themes in the breast cancer classification literature revolve around the development of histopathological and medical imaging datasets that serve as an important foundation for training machine learning models. The most cited articles focus on: Providing high-quality histopathological and ultrasound datasets, which are the main ingredients for building AI-based automated classification models. Classification and prognostic markers, indicating a great interest in utilizing medical data for more accurate cancer prediction and detection. The direction of research is strongly influenced by the availability of open datasets and the integration of digital image analysis with machine learning methods such as Random Forest and Extreme Learning Machine. These findings emphasize the importance of input data quality and information structure in developing efficient and accurate AI-based cancer diagnosis systems.

RESULT AND DISCUSSION

Major Classification Approaches in Breast Cancer Research

This study reveals the dominant classification approaches in the literature related to breast cancer classification, especially those involving the Random Forest (RF) and Extreme Learning Machine (ELM) algorithms. RF stands out as a widely used ensemble method due to its ability to produce stable and accurate classifications in complex data environments (Aroef et al., 2020; Quist et al., 2021). RF combines multiple independently trained decision trees, then combines their prediction results through majority voting, which increases accuracy and reduces the risk of overfitting (Octaviani & Rustam, 2019). Based on the analyzed datasets, RF is widely applied in the classification of benign and malignant cells, as well as in feature processing in histopathological and mammographic images. The use of ensemble methods such as RF allows handling of noise and outliers that are often found in medical data. A study by Gupta et al. (2020) also showed that RF showed consistently high performance on medical image datasets, even when dealing with high-dimensional data. Meanwhile, ELM emerged as an alternative that offers high training speed and computational efficiency (Zhang et al., 2021). ELM is designed as a single hidden layer artificial neural network whose weights are randomly initialized, and its training is completed by mathematical solutions without iterations, making it very computationally efficient (Zhang & Chen, 2020). ELM is well suited for the development of real-time classification systems and has been used in several studies to process breast image data with minimal computational time.

However, ELM also has limitations, such as sensitivity to the number of hidden neurons and initial weight distribution, which can affect performance stability. Several studies have proposed the integration of ELM with ensemble methods or kernel optimization to improve model generalization and stability, but such studies are still relatively limited in the breast cancer domain. Overall, both RF and ELM make important contributions to the development of breast cancer classification systems, but each has its own advantages and challenges that must be considered in clinical applications. Further research is highly recommended to explore strategies for combining the two or compare performance across public datasets with varying class and feature distribution conditions.

Cost Sensitivity and Unbalanced Data Handling

One of the major issues in breast cancer classification is the imbalanced class distribution, where benign cancer data is much more dominant than malignant cancer data. Many studies in this dataset have shown that the application of cost-sensitive algorithms, especially on RF, helps improve the detection of minority (malignant) classes. By giving a larger penalty to false negatives, the model becomes more sensitive in recognizing malignant cancer cases that have significant clinical impact (Sharma & Sunkaria, 2022). Commonly used strategies include cost-sensitive RF, reweighting, and the use of modified loss functions. However, only a few studies explicitly incorporate cost-sensitive approaches into the ELM architecture, reflecting an underexplored research opportunity. This could be due to the limited ability of ELM to handle internal weighting or the absence of a standard framework for cost-sensitive ELM, so future research can focus on developing a version of ELM that explicitly considers minority class penalties (Gupta et al., 2020).

Computational Complexity and Model Efficiency

The computational efficiency aspect is also an important highlight in this study. RF is known to require higher training time as the number of trees and features analyzed increases. The algorithmic complexity of RF is around $O(n \times m \times \log m)$, where n is the number of trees and m is the number of features. Although very accurate, RF tends to be unsuitable for real-time applications or systems with hardware constraints. Meanwhile, ELM shows advantages in training speed and resource usage because it does not involve an iterative process in its training. ELM can be solved by matrix computation directly, making it ideal for deployment in low-power systems (Zhang et al., 2021). However, there is a trade-off between accuracy and efficiency, where several studies have shown a decrease in ELM performance on highly imbalanced or complex datasets (Zhang & Chen, 2020). Therefore, hyperparameter optimization and hybrid techniques have become strategies that are increasingly being proposed to combine the advantages of RF (high accuracy) and ELM (high efficiency). Future studies can also evaluate computational efficiency based on inference time, memory consumption, and scalability metrics to strengthen model selection recommendations in real clinical contexts.

Performance Metrics in Breast Cancer Classification and Its Impact

With Python, we analyzed the performance metrics in breast cancer classification and its impact as written below:

```
# Import the required library
import pandas as pd

# Assume the data is already loaded into a DataFrame named df_combined
# Combine title and abstract for text analysis
df_combined['combined_text'] = df_combined['Title'].fillna('') + ' ' + df_combined['Abstract'].fillna('')

# Filter articles containing keywords related to engagement metrics
keyword_filter = r'engagement|metrics|performance|sensitivity|accuracy|specificity|recall|AUC|training time|efficiency'
engagement_content = df_combined[df_combined['combined_text'].str.contains(keyword_filter, case=False, na=False)]

# Sort by highest citation count
top_engagement_research = engagement_content.sort_values(by='Cites', ascending=False).head(10)

# Display the top articles on engagement metrics
print("Top research findings on engagement metrics in breast cancer classification:")
print(top_engagement_research[['Title', 'Year', 'Cites']])
```

Table 2. Performance Metrics Categories in Breast Cancer Classification and Their Impact

Category	Metric	Clinical/Research Impact	Value for AI Systems
Model Validation	Accuracy	Reflects overall performance of the classification model	General performance evaluation
Model Validation	AUC-ROC	Assesses trade-off between sensitivity and specificity	Classification reliability

JDBIM (Journal of Digital Business and Innovation Management)
Volume 4, No. 2, December 2025
E-ISSN 2692-3898
Page 143-162

Class Imbalance Handling	Sensitivity (Recall)	Correct detection of malignant cancer cases	Minimization of False Negatives
Class Imbalance Handling	Specificity	Accurate identification of benign (non-cancer) cases	Minimization of False Positives
Model Efficiency	Training Time	Determines feasibility for real-time implementation	Computational efficiency
Model Efficiency	Algorithmic Complexity	Measures computational burden on resource-limited systems	System optimization
Model Stability	Output Variance Across Folds	Indicates model resilience to data fluctuation	Replicability and model generalization
Interpretability	Feature Importance Score	Identifies most relevant clinical features	Transparency & explainability
Dataset Reliability	Dataset Size & Class Balance	Affects training and generalization of the model	Long-term model validity
Clinical Relevance	F1-Score	Balances precision and recall in medical context	Overall classification effectiveness

The results presented in Table 2 show that the evaluation metrics for breast cancer classification models can be grouped into several main categories, namely: model validation, handling class imbalance, model efficiency, stability, interpretability, dataset reliability, and clinical relevance. Each metric plays an important role in assessing various aspects of an AI-based classification system. For example, accuracy and AUC-ROC are often used to describe the overall performance of a model and its ability to distinguish benign from malignant cases. In the context of handling imbalanced data, metrics such as sensitivity and specificity become very crucial because they ensure that the model is able to correctly identify malignant cancers without many false positives. In terms of practical implementation, training time and algorithmic complexity determine whether a model is feasible to be used in real-world situations that have limited resources or require rapid real-time responses. In addition, feature importance score contributes to model interpretability, which is crucial in clinical settings where transparency and trust in the system are essential. Metrics related to dataset characteristics such as data size and balance also affect the training quality and generalization ability of the model. Finally, F1-score is an indicator that balances precision and recall, which is particularly relevant for medical cases where misclassification can have a significant impact. These findings underscore the importance of using comprehensive and contextual evaluation metrics in the development of machine learning-based diagnostic systems.

Practical Implications and Further Research Directions

The results of this systematic review have several important practical implications for the development of breast cancer classification systems in the future. First, the application of algorithms such as Random Forest and Extreme Learning Machine can be adjusted to the needs of AI-based diagnostic systems in hospitals, especially in the context of histopathology image classification. RF can be utilized for models that require high interpretability and accuracy, while ELM is suitable for real-time scenarios that require computational speed. Second, the importance of integrating cost-sensitive approaches in the classification of malignant cancers signals that the developed system must consider the weight of misdiagnosis differently. AI-based systems should not only pursue high accuracy in general, but also minimize the risk of false negatives. Third, future research needs to adopt a hybrid approach that combines the strengths of RF and ELM, and integrates automatic parameter optimization and ensemble techniques. In addition, there is a great opportunity to develop edge computing-based systems so that diagnosis can be carried out locally with high efficiency, especially in areas with limited access to advanced health facilities. With a systematic approach, the results of this study are expected to be a reference for the development of breast cancer classification systems that are more adaptive, efficient, and ready to be applied in the real world.

JDBIM (Journal of Digital Business and Innovation Management)
Volume 4, No. 2, December 2025
E-ISSN 2692-3898
Page 143-162

CONCLUSION

This systematic review reveals the comparative performance and strategic relevance of Random Forest (RF) and Extreme Learning Machine (ELM) algorithms in breast cancer classification, especially in terms of cost sensitivity and computational complexity. RF is shown to provide consistently high accuracy and has good robustness to high-dimensional data, making it suitable for use in complex diagnostic environments requiring interpretability. On the other hand, ELM offers advantages in training speed and computational efficiency, which are particularly useful for real-time diagnostic systems and resource-constrained devices, although its performance may fluctuate depending on parameter sensitivity. This review also highlights the importance of cost-sensitive learning strategies in addressing the problem of data imbalance that is common in breast cancer studies. Several studies have shown that penalizing false negatives can significantly improve model sensitivity, which is crucial in clinical contexts. However, the application of this approach in ELM architecture is still very limited, opening up great opportunities for further development.

In terms of computational efficiency, RF tends to require more resources due to its ensemble nature, while ELM is a promising lightweight alternative for systems with low latency requirements. However, there is a trade-off between accuracy and speed, highlighting the importance of exploring hybrid models and optimizing parameters. Limitations of this review include the reliance on secondary literature and published data, which may not reflect recent developments that have not been widely published. In addition, the lack of standardization of datasets used in various studies may affect the reliability of model comparisons. For future research, it is recommended to develop hybrid models that combine the strengths of RF and ELM, design cost-sensitive ELM architectures, and test computational complexity based on actual inference time. The application of explainable AI (XAI) principles is also important to increase trust and adoption in clinical settings. Thus, this review provides a strong conceptual foundation for the development of AI-based breast cancer diagnostic systems that are more adaptive, efficient, and tailored to the needs of medical practice in various healthcare settings.

REFERENCES

Aroef, C., Rivan, Y., & Rustam, Z. (2020). Comparing random forest and support vector machine for breast cancer classification. TELKOMNIKA

Telecommunication, Computing, Electronics and Control, 18(3), 1494–1501. https://doi.org/10.12928/telkomnika.v18i3.14658

Chen, J., & Su, X. (2022). Measuring the effectiveness of live selling strategies. Decision Support Systems, 159, 113712. https://doi.org/10.1016/j.dss.2022.113712

Chen, J., & Wu, C. (2022). The role of interactivity in live e-commerce: Implications for customer engagement and purchase intention. Journal of Business Research, 145, 222–234. https://doi.org/10.1016/j.jbusres.2022.03.010

Gupta, D., Khanna, A., & Koundal, D. (2020). Breast cancer prediction using computational intelligence techniques: A comparative study. Computers in Biology and Medicine, 125, 104013. https://doi.org/10.1016/j.compbiomed.2020.104013

Harzing, A. W. (2007). Publish or Perish. https://harzing.com/resources/publish-or-perish

Li, Y., & Guo, L. (2022). The effect of social presence in live-streaming on consumer behavior. International Journal of E-commerce Research, 18(3), 145–160. https://doi.org/10.1007/s10660-021-09442-z

Liang, C., & He, P. (2021). Consumer behavior in e-commerce live streaming: Insights from China. Journal of Retailing and Consumer Services, 59, 102883. https://doi.org/10.1016/j.jretconser.2020.102883

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Medicine, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Octaviani, T. L., & Rustam, Z. (2019). Random forest for breast cancer prediction. AIP Conference Proceedings, 2194(1), 020060. https://doi.org/10.1063/1.5139820

Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Quist, J., Taylor, L., Staaf, J., & Grigoriadis, A. (2021). Random forest modelling of high-dimensional microarray data for breast cancer subtype classification. Cancers, 13(5), 991. https://doi.org/10.3390/cancers13050991

Sharma, R., & Sunkaria, R. K. (2022). A cost-sensitive random forest classifier for breast cancer detection using synthetic data augmentation. Expert Systems with Applications, 192, 116381. https://doi.org/10.1016/j.eswa.2021.116381

Zhang, W., & Chen, X. (2020). A hybrid extreme learning machine for breast cancer diagnosis. Applied Soft Computing, 86, 105778. https://doi.org/10.1016/j.asoc.2019.105778

Zhang, Y., Wu, L., & Wang, J. (2021). Breast cancer classification using extreme learning machine with improved kernel function. Knowledge-Based Systems, 221, 106911. https://doi.org/10.1016/j.knosys.2021.106911

Zhou, Y., & Wang, H. (2020). The impact of social presence on purchase intention in live commerce. Journal of Business and Industrial Marketing, 35(8), 1225–1238. https://doi.org/10.1108/JBIM-07-2020-0325