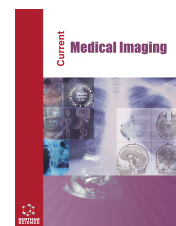




Current Medical Imaging

Content list available at: <https://benthamscience.com/journals/cmimr>



RESEARCH ARTICLE

Application of Tuning-ensemble N-Best in Auto-Sklearn for Mammographic Radiomic Analysis for Breast Cancer Prediction

Faikah Awang Ismail^{1,2}, Muhammad Khalis Abdul Karim^{2,3,*}, Siti Izzatul Akma Zaidon³ and Kaltham Abdulwahid Noor⁴

¹School of Biology, Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Negeri Sembilan, Kampus Kuala Pilah, 72000, Malaysia

²Department of Physics, Faculty of Science, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia

³Institute for Mathematical Research (INSPERM), Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia

⁴Dubai Health, Radiology Department, Rashid Hospital, Dubai 00971, United Arab Emirates

Abstract:

Introduction:

Breast cancer is a major cause of mortality among women globally. While mammography remains the gold standard for detection, its interpretation is often limited by radiologist variability and the challenge of differentiating benign and malignant lesions. The study explores the use of Auto-Sklearn, an automated machine learning (AutoML) framework, for breast tumor classification based on mammographic radiomic features.

Methods:

244 mammographic images were enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) and segmented with Active Contour Method (ACM). Thirty-seven radiomic features, including first-order statistics, Gray-Level Co-occurrence Matrix (GLCM) texture and shape features were extracted and standardized. Auto-Sklearn was employed to automate model selection, hyperparameter tuning and ensemble construction. The dataset was divided into 80% training and 20% testing set.

Results:

The initial Auto-Sklearn model achieved an 88.71% accuracy on the training set and 55.10% on the testing sets. After the resampling strategy was applied, the accuracy for the training set and testing set increased to 95.26% and 76.16%, respectively. The Receiver Operating Curve and Area Under Curve (ROC-AUC) for the standard and resampling strategy of Auto-Sklearn were 0.660 and 0.840, outperforming conventional models, demonstrating its efficiency in automating radiomic classification tasks.

Discussion:

The findings underscore Auto-Sklearn's ability to automate and enhance tumor classification performance using handcrafted radiomic features. Limitations include dataset size and absence of clinical metadata.

Conclusion:

This study highlights the application of Auto-Sklearn as a scalable, automated and clinically relevant tool for breast cancer classification using mammographic radiomics.

Keywords: Breast cancer, Machine learning, Auto-Sklearn, Mammography, Radiomic analysis, ACM.

Article History

Received: April 22, 2025

Revised: May 14, 2025

Accepted: May 19, 2025

1. INTRODUCTION

Breast cancer is one of the most common causes related to cancer that affects women globally, with 2.3 million new cases reported, and with approximately 12% of all new cancer diagnoses globally [1]. Early detection significantly improves survival rates, and mammography is regarded as the gold standard for breast cancer screening and diagnosis. However,

the manual interpretation of mammography continues to pose challenges due to radiologist variability, subjectivity and the difficulty in differentiating between benign and malignant tumors [2, 3]. In addition to that, dense breast tissues can obscure tumor detection, leading to false negatives, while some benign cases may be misdiagnosed as malignant, resulting in unnecessary biopsies [4]. These limitations highlight the need

for artificial intelligence (AI) and machine learning (ML)-based approaches to improve precision and reduce diagnostic discrepancies [5, 6]. Recent deep learning models, including Capsule Networks and transformer-based segmentation, have demonstrated improvements in diagnostic accuracy and interpretability in disease detection through enhanced Region of Interest (ROI) extraction and feature classification [7, 8]. However, it requires high-end GPU and expertise to be implemented. User-friendly and efficient approaches are needed to make it accessible for wider application and clinical settings.

Radiomic feature extraction helps in quantitative analysis for medical images, which is crucial in diagnosing breast cancer. Denoising and feature enhancement applied before feature extraction help in increasing image quality while preserving structural integrity, especially when multiscale features are required [9 - 11]. However, some approaches needed expertise in coding, which might be a rarity in clinical settings. Thus, in this study, CLAHE and ACM methods, which are user-friendly, were utilized before feature extraction in this study to achieve optimal quality feature extraction.

Automated Machine Learning (AutoML) has gained attention as a powerful tool for automating model selection, feature extraction, and hyperparameter tuning. Auto-Sklearn, one of the AutoML frameworks, has demonstrated efficiency in optimizing machine learning (ML) models through Bayesian Optimization (BO), meta-learning, and ensemble techniques [12]. Auto-Sklearn enhances classification performance by automating the labor-intensive processes of ML model creation, hence reducing the need for human intervention, which makes it suitable for medical imaging applications. In comparative studies involving other imaging modalities such as brain tumor and Alzheimer's detection, deep learning models like ResNet, U-Net, attention mechanisms, and more advanced N-Beats have consistently outperformed conventional methods in feature extraction and lesion classification accuracy [7, 13 - 15]. These findings support the integration of AutoML with robust image preprocessing and segmentation methods to improve outcomes in radiomic breast cancer prediction studies. While previous studies have demonstrated promising advances in medical image classification using deep learning, these approaches primarily focus on manually designed pipelines that require expert knowledge for feature selection, model tuning, and architecture design, which might be challenging to implement in clinical settings. In contrast, this study introduces AutoML-based pipeline that automates not only model selection and hyperparameter tuning, but also ensemble construction, thereby making it more user-friendly for clinical applications.

Due to limitations of manual image processing and reporting, there is an urgent need for an advanced approach to increase the speed of breast cancer detection while optimizing the diagnosis quality *via* AI. In this study, Tuning-Ensemble N-

Best in Auto-Sklearn is explored for mammographic radiomic analysis in breast cancer predictions. The methods involve the extraction of radiomic features from mammographic images, including first-order, GLCM texture second order and shape-based features, followed by the application of Auto-Sklearn for tumor classification. Lastly, the performance of standard Auto-Sklearn models with resampling strategies will be evaluated. By leveraging AutoML for mammographic analysis, this study aims to contribute to the development of more efficient, accurate, and automated diagnostic tools for breast cancer detection.

2. MATERIALS AND METHODS

2.1. Data sources

The sample used in this study was 244 mammogram images with Craniocaudal (CC) and Mediolateral Oblique (MLO) views, extracted from two databases, Curated Breast Imaging Subset of Digital Database for Screening Mammography and Digital Database for Screening Mammography (CBIS-DDSM) <https://www.cancerimagingarchive.net/collection/cbis-ddsm/>.

The samples were categorized as benign or malignant cases based on pathologically confirmed diagnoses made by the experts. After all images were converted into DICOM format, the data were preprocessed and analyzed in MATLAB version R2022a for image enhancement *via* CLAHE and image segmentation *via* ACM. MATLAB programming was used for feature extraction. Later the employment of Auto-Sklearn framework was implemented in Python version 3.9 with Sci-kit Learn for automated model selection, hyperparameter tuning, and model performance evaluation. The study was conducted on a standard desktop computer equipped with an Intel Core i5 (10th generation) processor, 16 GB RAM, and a 512 GB SD.

2.2. Image Enhancement and Image Segmentation

Fig. (1) illustrates the workflow of the study. The CLAHE was employed to enhance the image contrast, where it utilized the adaptive histogram equalization approach to separate the image into specific parts called "tiles" and performed histogram equalization separately on each tile. The method avoids amplification noise in the image by constraining the amplification contrast in each tile. Each image dataset's contrast of mammogram pictures was enhanced using CLAHE. In this study, the clip limit was set to 0.02 as the most suitable level of contrast, and the number of tiles was set to the default [8 x, 8]. Fig. (2) shows the contrast difference between the images prior to and after the application of CLAHE.

After that, the images were enhanced *via* MATLAB R2022a using the "*adapthisteq*" function, the images of the dataset were compared with the images of segmentation that were already segmented by radiologists with manual segmentation. Then, the procedure of segmenting an image starts with the intended ROI on the enhanced mammography images [16, 17]. Later, the ACM technique was applied for the already enhanced images to refine the boundaries of the ROI by iteratively adjusting contour points to fit the tumour shape [18].

* Address correspondence to this author at the Department of Physics, Faculty of Science, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia and Institute for Mathematical Research (INSPERM), Universiti Putra Malaysia, 43400, UPM Serdang, Selangor Malaysia; E-mail: mkhalis@upm.edu.my

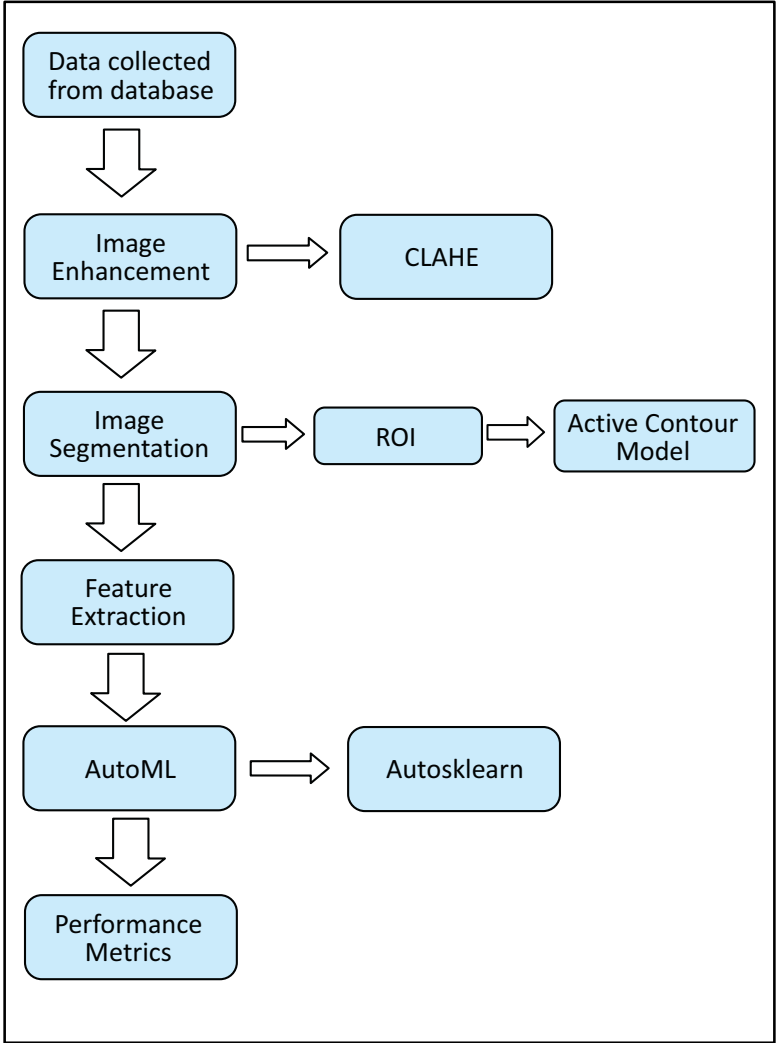


Fig. (1). Workflow for study.

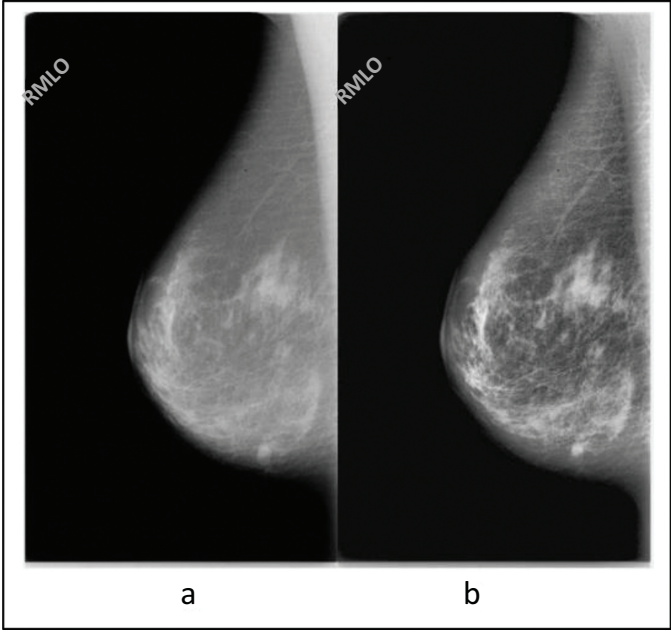


Fig. (2). Mammogram images in MLO view, a) the image without CLAHE while b) the image enhanced by CLAHE.

In this study, a semi-automatic segmentation approach was used where radiologists initially marked potential tumour regions and ACM refined the segmentation. The ACM energy function consists of three components: elasticity, rigidity, and image gradient attraction, formulated as (1):

$$E_{contour} = \int (\alpha |v'(s)|^2 + \beta |v''(s)|^2 + \gamma |\nabla I(v(s))|^2) ds \quad (1)$$

Where $v(s)$ represent contour at position (s) , α controls the elasticity of the contour, β maintains the rigidity and γ ensures the contour moves towards the image gradient. The number of iterations for ACM was set to 150, ensuring precise tumour boundary detection Fig. (3). Illustrates mammography images that are being segmented using ACM.

2.3. Feature Extraction

Once the image preprocessing was completed, radiomic features were extracted to quantify tumor characteristics. 37 features were extracted from each mammogram. These features were divided into three categories: Six first-order statistical, 21 GLCM texture second-order, and ten shape features. First-order statistical features encompassed basic intensity descriptors such as mean, variance, skewness, kurtosis, energy, and entropy, which describe the pixel intensity distribution in

an image [19].

Second-order texture features using GLCM, which include contrast, correlation, homogeneity, energy, and dissimilarity, provide information about spatial relationships between pixel intensities [20]. Lastly, shape-based features, including area, perimeter, eccentricity, convexity, and solidity, were retrieved to examine the geometric aspects of tumours [21 - 23]. These features were subsequently normalized and standardized for ML classification. Standardization ensured that all features are on the same scale, preventing bias towards feature values in a range of 0 to 1. Several features are mathematically defined in Eq. (2) as follows:

$$Contrast = \sum_{i,j} (i - j)^2 P(i, j) \quad (2)$$

where $P(i, j)$ is the probability of the pixel intensities i and j occurring together. Furthermore, homogeneity Eq. (3):

$$Homogeneity = \sum_{i,j} \frac{P(i, j)}{1 + |i - j|} \quad (3)$$

where values of 1, indicate smoother texture. And finally, the entropy Eq. (4)

$$Entropy = - \sum_{i,j} P(i, j) \log P(i, j) \quad (4)$$

where, the higher value indicates higher complexity.



Fig. (3). Segmented area in mammography images ACM.

2.4. Model Training using Auto-Sklearn

The Auto-Sklearn model used in this study was designed to automate the selection, tuning and training of ML models for breast cancer classification. This AutoML utilizes BO, meta-learning and ensemble learning techniques to enhance model performance without requiring manual intervention [17]. It identifies the best-performing classification model based on radiomic features extracted from mammographic images and optimizes hyperparameters for improved accuracy and robustness of the model.

The training process began by uploading the extracted radiomic features obtained from segmented mammogram images. The dataset was later divided into 80% for training and 20% for testing subsets, using stratified sampling to preserve the original class distribution between benign and malignant cases. This approach ensures that both subsets reflect the same proportion of cases as the overall dataset, thereby minimizing potential sampling bias. The Auto-Sklearn library was initialized, and the “fit” function was used to train multiple ML models on the dataset. Auto-Sklearn automatically explored a variety of classification algorithms, including Support Vector Machines (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), to select the best-performing model using cross-validation scores. For SVM, the hyperparameter “c” was sampled from a log-uniform distribution with the range $1e-3, 1e3$. For Random Forest, the number of estimators ($n_estimators$) was selected from the discrete range of 10 to 1000 and lastly, for XGBoost, learning rates were examined within the range of 0.01 to 1.0. The BO seeks the optimal set hyperparameters, θ^* by maximizing an objective function (5):

$$\theta^* = \arg \max_{\theta \in \Theta} f(\theta) \quad (5)$$

where Θ represent hyperparameter space, and $f(\theta)$ is the objective function.

2.5. Model Performance Evaluation

The model performance was evaluated *via* the value of accuracy, precision, recall (sensitivity), specificity, F1-score, and the AUC-ROC. Accuracy is defined as the ratio of correct predictions to total predictions. Precision measured the proportion of correctly identified malignant cases among all predicted malignant cases, while recall (sensitivity) quantified the model’s ability to detect actual malignant cases. Below Eq. (6) are some mathematical definitions of model evaluation:

$$TPR = \frac{TP}{TP+FN} \quad (6)$$

where TPR is the True Positive Rate, TP, true positive, and FN is false negative. Furthermore, the False Positive Rate (FPR) is defined as Eq. (7):

$$FPR = \frac{FP}{FP+TN} \quad (7)$$

where FP and TN are known as False Positive and True Negative, respectively. Meanwhile, the AUC-ROC, mathematically defined as Eq (8):

$$AUC = \int_0^1 TPR(FPR)d(FPR) \quad (8)$$

The specificity score assessed the model’s ability to accurately identify benign cases, thereby maintaining a balanced classification performance. The F1-score, defined as the harmonic mean of precision and recall, was computed to evaluate the balance between false positive and false negative cases.

The ROC curve was plotted to evaluate the true positive rate (sensitivity) about the false positive rate (1-specificity) at varying classification thresholds. The AUC-ROC score offers a comprehensive assessment of the model’s capacity to differentiate between benign and malignant tumours, with values approaching 1.0 signifying enhanced classification efficacy [24].

3. RESULTS AND DISCUSSIONS

In this section, the study’s findings were highlighted. The first-order statistical features were found to contribute significantly to model performance, where entropy and variance showed the highest correlation with malignancy, consistent with previous studies that indicated the role of texture irregularity in tumor characterization [25, 26]. Second-order texture features in GLCM, such as contrast and heterogeneity, were found to be strong predictors of malignancy. These features delineate spatial interactions between pixels and offer essential insights into tumor heterogeneity, which is also known as an indicator of cancer aggressiveness [27, 28]. Shape-based features such as eccentricity and convexity further supported classification by capturing geometric differences between benign and malignant tumors, as the latter typically possess uneven shapes with serrated edges, whereas the former generally have smooth, well-defined margins [29, 30].

The performance gap observed between the training accuracy (88.71%) and testing accuracy (55.10%) in the initial Auto-Sklearn model suggests overfitting, where the model performs well on the training data but fails to generalize to unseen data [12]. This anomaly is likely attributed to several factors, notably the limited dataset (244 samples) and high dimensionality of radiomic features (37 features), which may introduce feature redundancy.

To enhance the model performance, the Tuning-Ensemble N-Best approach was implemented, where the top-performing models were combined into an ensemble ML. Additionally, BO was applied to fine-tune hyperparameters, systematically exploring parameter configurations to optimize classification performance. This technique is useful in reducing the number of evaluations required to reach an optimal solution, making Auto-Sklearn computationally efficient. The whole process was repeated 15 times to mitigate the impact of randomness in the model selection and training process, resulting in more consistent and dependable outcomes. This can be seen in Fig. (4), where the accuracy score of the model increases after the after-resampling strategy has been applied and repeated 15 times.

Before the resampling, the model tended to favor the majority class (benign cases), leading to poor sensitivity in detecting malignancies. Following resampling, the training accuracy increased to 95.26%, while the testing set accuracy

improved to 76.16%, where the recall for malignant cases increased by approximately 18%, improving the model’s ability to correctly identify tumours. This finding aligns with previous studies that emphasize the importance of balancing training datasets to prevent bias in ML models [31, 32].

The enhanced precision and F1-score upon resampling demonstrate that Auto-Sklearn efficiently minimizes false positives and false negatives, making it a promising tool for radiologists in decision-making. After the resampling, the AUC-ROC also improved from 0.660 to 0.840, indicating an enhanced capacity to differentiate between benign and malignant cases for the model (Table 1 and Fig. 5). From a clinical perspective, this level of discriminative performance is meaningful, as it suggests that the model could potentially

assist radiologists in improving early diagnostic accuracy for breast cancer classification. A rise in AUC-ROC signifies a more dependable classification model, with a value approaching 1.0, indicating superior discrimination between two groups [33]. A high AUC-ROC indicates that the model maintains a favorable balance between sensitivity and specificity across different decision thresholds, which is essential in breast cancer screening, where both false negatives (missed malignancies) and false positives (unnecessary biopsies) carry significant consequences. Fig. (5) shows the ROC-AUC value for lesion classification increases as compared to the standard processing after application of the resampling strategy, where the x-axis represents False Negative Rate and the y-axis represents True Positive Rate.

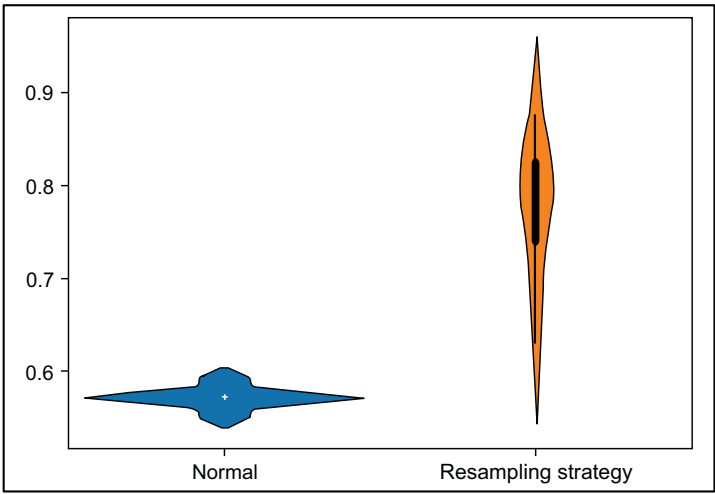


Fig. (4). Range of accuracy score for Auto-Sklearn after 15 repetitions.

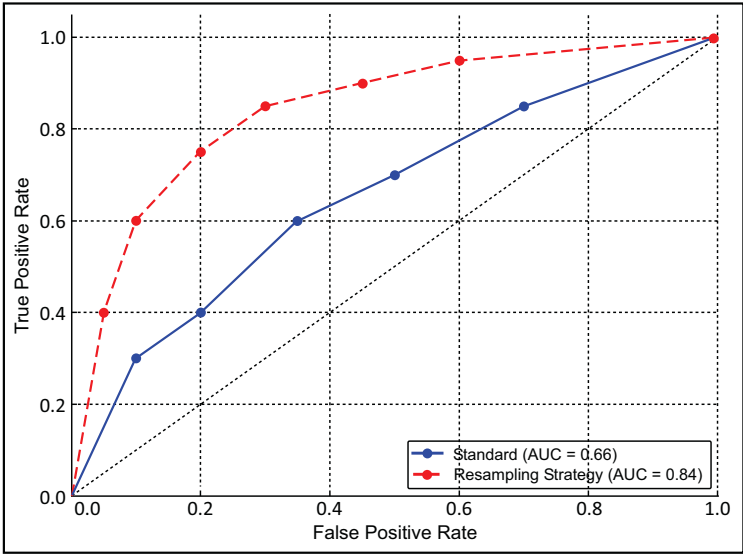


Fig. (5). The ROC-AUC graph, blue line indicates performance of standard processing and red line indicates the performance after the application of the resampling strategy for lesion classification.

Table 1. Performance metric value between standard and after resampling strategy.

Evaluation Metric	Standard	Resampling Strategy
Accuracy	0.551	0.792
Precision	0.609	0.704
Recall	0.538	0.760
ROC AUC	0.660	0.840

A comparative analysis indicates that Auto-Sklearn surpassed conventional models in terms of adaptability and automated hyperparameter tuning. The SVM model achieved an accuracy of 72.3%, RF scored 74.1% and XGBoost obtained 75.6%, whereas Auto-Sklearn with resampling reached 76.16%, underscoring the benefits of automated model selection and optimization. This is consistent with findings from previous studies, where AutoML frameworks demonstrated improved accuracy and efficiency relative to manually optimized ML models [12]. The ability of Auto-Sklearn to automatically select the best-performing algorithms and optimize hyperparameters significantly reduces the time and expertise required for model development, making it effective for medical imaging applications in clinical settings.

The computational efficiency of Auto-Sklearn was also evaluated. The total computation time for Auto-Sklearn training and optimization was significantly lower compared to manual hyperparameter tuning in conventional models. In our study, Auto-Sklearn was able to complete the optimization process within 30-40 minutes, as it enabled fewer steps for ML processing for our dataset; however, it might not be applicable for studies involving larger datasets and more complicated processing. Previous studies demonstrate the necessity for more powerful GPUs and extended training times for dataset processing, particularly when dealing with huge datasets that encompass complex architecture and multi-modal data [14, 34].

While Auto-Sklearn includes built-in resampling and ensemble strategies to mitigate overfitting, it might be insufficient in small-scale radiomic studies. To further lower the risk of overfitting, future work could incorporate explicit feature selection techniques such as Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA), and regularization techniques.

Despite the promising results, the model exhibited several misclassifications that warrant further analysis. Specifically, false-positive cases, which are often associated with benign tumors in dense breast tissue, may appear irregular or heterogeneous in texture, leading the model to misclassify them as malignant. Conversely, some malignant tumors with relatively smooth boundaries were predicted as benign, likely due to their lower texture complexity or homogeneous shape. However, by relying solely on handcrafted radiomic features, some subtle morphological or contextual cues present in mammographic images might be undetected.

Additionally, a small dataset (n=244) and absence of complementary information such as clinical metadata may reduce the model's ability to differentiate borderline cases. The relevance of this study's findings can be strengthened by external validation, as it might help to train the models in the

clinical setting. These challenges indicate the need to integrate deep learning approaches such as Convolutional Neural Networks (CNNs), with AutoML, which could further improve the classification performance. CNNs excellence in feature extraction may boost AutoML by enhancing spatial feature representation. The hybrid methods, such as ResNet and DenseNet could play a role as feature extractors *via* transfer learning, where deep features derived from mammographic images become an input into the AutoML pipeline for model selection and optimization. Previous studies suggested transformer-based model approaches, especially when the diagnosis involved multi-layer image processing and multi-imaging modalities, as they provide higher model performance compared to conventional deep learning approaches [7, 14, 34]. This approach can be considered in the future for further studies, especially in breast cancer diagnosis.

4. STUDY LIMITATIONS

While the findings of this study are promising and demonstrate the potential of Auto-Sklearn for radiomic-based breast cancer classification, several limitations must be acknowledged. The sample size (n=244), though acceptable for an exploratory study, may restrict the generalizability of the model to larger, more diverse clinical populations. The utilization of publicly accessible CBIS-DDSM data, despite its standardization, may not entirely capture the diversity in actual clinical settings, thereby leading to selection bias. Moreover, the reliance on handcrafted radiomic features may not capture the full complexity of tumor heterogeneity, especially in cases involving dense breast tissue or subtle malignancies. As this is an exploratory study, the absence of external validation and the exclusion of clinical metadata may have reduced the model's ability to handle borderline or ambiguous cases. Nevertheless, the study provides a solid foundation for the use of AutoML in breast cancer diagnosis.

CONCLUSION

This study demonstrates the feasibility and effectiveness of integrating Auto-Sklearn, with radiomic feature analysis for the automated classification of breast tumors using mammographic images. By incorporating user-friendly preprocessing methods such as CLAHE and ACM for image enhancement and segmentation, the study successfully extracted 37 radiomic features and achieved a post-resampling testing accuracy of 76.16% with an AUC-ROC of 0.840. These findings highlight AutoML's capacity to automate model development and improve reproducibility in medical imaging workflows, particularly in clinical settings with limited technical expertise. The contribution of this study lies in its development of a user-friendly, scalable, and effective diagnostic pipeline that minimizes manual intervention while preserving clinical

relevance. Future studies should focus on incorporating larger, multi-institutional datasets and external validation to enhance model generalizability. Moreover, combining AutoML with deep learning approaches such as CNNs and transformer-based architectures may enable the integration of handcrafted and deep features, further improving classification performance in real clinical settings.

AUTHORS' CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: M.K.A.K.: Study conception and design; S.I.A.Z.: Methodology; K.A.N.: Investigation; F.A.I.: Data collection; All authors reviewed the results and approved the final version of the manuscript.

LIST OF ABBREVIATIONS

ACM	= Active Contour Method
AI	= Artificial Intelligent
AUC	= Area Under Curve
AutoML	= Automated Machine Learning
BO	= Bayesian Optimization
CBIS	= Curated Breast Imaging Subset
CC	= Craniocaudal
CNN	= Convolutional Neural Networks
CLAHE	= Contrast Limited Adaptive Histogram Equalization
GLCM	= Gray Level Co-occurrence Matrix
DDSM	= Digital Database for Screening Mammography
ML	= Machine Learning
MLO	= Mediolateral Oblique
RF	= Random Forest
ROC	= Receiver Operating Curve
SVM	= Support Vector Machines
XGBoost	= Extreme Gradient Boosting

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIAL

All data generated or analyzed during this study are included in this published article.

FUNDING

The work has been supported by Universiti Putra Malaysia, Malaysia under vot GPI/UPM/9762000.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Department of Physics and INSPM, Universiti Putra Malaysia, for technical and computational support. Appreciation is also extended to the School of Biology, UiTM Cawangan Negeri Sembilan, and the Radiology Department, Rashid Hospital, Dubai for their contribution.

REFERENCES

- [1] Sung H, Ferlay J, Siegel RL, *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71(3): 209-49. [http://dx.doi.org/10.3322/caac.21660] [PMID: 33538338]
- [2] Lohmann P, Bousabarah K, Hoevels M, Treuer H. Radiomics in radiation oncology—basics, methods, and limitations. *Strahlenther Onkol* 2020; 196(10): 848-55. [http://dx.doi.org/10.1007/s00066-020-01663-3] [PMID: 32647917]
- [3] Radzi SF, Karim MKA, Saripan MI, Abd Rahman MA, Osman NH, Dalah EZ. Impact of image contrast enhancement on stability of radiomics feature quantification on a 2d mammogram radiograph. *IEEE Access* 8: 127720. [http://dx.doi.org/10.1109/ACCESS.2020.3008927]
- [4] Li H, Zhuang S, Li D, Zhao J, Ma Y. Benign and malignant classification of mammogram images based on deep learning. *Biomed Signal Process Control* 2019; 51: 347-54. [http://dx.doi.org/10.1016/j.bspc.2019.02.017]
- [5] Lee JY, Lee K, Seo BK, *et al.* Radiomic machine learning for predicting prognostic biomarkers and molecular subtypes of breast cancer using tumor heterogeneity and angiogenesis properties on MRI. *Eur Radiol* 2022; 32(1): 650-60. [http://dx.doi.org/10.1007/s00330-021-08146-8] [PMID: 34226990]
- [6] Sheng W, Xia S, Wang Y, *et al.* Invasive ductal breast cancer molecular subtype prediction by MRI radiomic and clinical features based on machine learning. *Front Oncol* 2022; 12: 964605. [http://dx.doi.org/10.3389/fonc.2022.964605] [PMID: 36172153]
- [7] Perumal K, Mahendran RK, Ahmad Khan A, Kadry S. Tri-M2MT: Multi-modalities based effective acute bilirubin encephalopathy diagnosis through multi-transformer using neonatal Magnetic Resonance Imaging. *CAAI Trans Intell Technol* 2025; 10: cit2.12409. [http://dx.doi.org/10.1049/cit2.12409]
- [8] Alhussen A, Anul Haq M, Ahmad Khan A, Mahendran RK, Kadry S. XAI-RACapsNet: Relevance aware capsule network-based breast cancer detection using mammography images *via* explainability O-net ROI segmentation. *Expert Syst Appl* 2025; 261: 125461. [http://dx.doi.org/10.1016/j.eswa.2024.125461]
- [9] Agrawal T, Choudhary P, Shankar A, Singh P, Diwakar M. MultiFeNet : Multi-scale feature scaling in deep neural network for the brain tumour classification in MRI images. *Int J Imaging Syst Technol* 2024; 34(1): e22956. [http://dx.doi.org/10.1002/ima.22956]
- [10] Diwakar M, Singh P, Garg D. Edge-guided filtering based CT image denoising using fractional order total variation. *Biomed Signal Process Control* 2024; 92: 106072. [http://dx.doi.org/10.1016/j.bspc.2024.106072]
- [11] Diwakar M, Kumar P, Singh P, Tripathi A, Singh L. An efficient reversible data hiding using SVD over a novel weighted iterative anisotropic total variation based denoised medical images. *Biomed Signal Process Control* 2023; 82: 104563. [http://dx.doi.org/10.1016/j.bspc.2022.104563]
- [12] Feurer M, Klein A, Eggensperger K, Springenberg JT, Blum M, Hutter F. Auto-sklearn: Efficient and robust automated machine learning. In: Hutter Frank, Kotthoff Lars, Vanschoren Joaquin, Eds. *Automated Machine Learning: Methods, Systems, Challenges*. Cham, Switzerland: Springer 2019; pp. 113-34. [http://dx.doi.org/10.1007/978-3-030-05318-5_6]
- [13] Gangadharan SMP, Dharani M, Thapliyal N, Yamsani N, Singh J, Singh P. Comparative analysis of Deep Learning-based brain tumor prediction models using MRI scan. 2023 Conference on Innovative

- Sustainable Computational Technologies (3rd International CISCT). Dehradun, India, 08-09 September 2023, pp. 1-6
[http://dx.doi.org/10.1109/CISCT57197.2023.10351227]
- [14] Khan AA, Madendran RK, Thirunavukkarasu U, Faheem M. *D²PAM*: Epileptic seizures prediction using adversarial deep dual patch attention mechanism. *CAAI Trans Intell Technol* 2023; 8(3): 755-69.
[http://dx.doi.org/10.1049/cit2.12261]
- [15] Kaur B, Rakhra M, Sharma N, *et al.* N-Beats architecture for explainable forecasting of multi-dimensional poultry data. *PLoS One* 2025; 20(4): e0320979.
[http://dx.doi.org/10.1371/journal.pone.0320979] [PMID: 40273069]
- [16] Bettinelli A, Marturano F, Avanzo M, *et al.* A novel benchmarking approach to assess the agreement among radiomic tools. *Radiology* 2022; 303(3): 533-41.
[http://dx.doi.org/10.1148/radiol.211604] [PMID: 35230182]
- [17] Mahmood T, Saba T, Rehman A, Alamri FS. Harnessing the power of radiomics and deep learning for improved breast cancer diagnosis with multiparametric breast mammography. *Expert Syst Appl* 2024; 249: 123747.
[http://dx.doi.org/10.1016/j.eswa.2024.123747]
- [18] Alshamrani K, Alshamrani HA, Alqahtani FF, Almutairi BS. Enhancement of mammographic images using histogram-based techniques for their classification using CNN. *Sensors (Basel)* 2022; 23(1): 235.
[http://dx.doi.org/10.3390/s23010235] [PMID: 36616832]
- [19] Kim D, Jensen LJ, Elgeti T, Steffen IG, Hamm B, Nagel SN. Radiomics for everyone: A new tool simplifies creating parametric maps for the visualization and quantification of radiomics features. *Tomography* 2021; 7(3): 477-87.
[http://dx.doi.org/10.3390/tomography7030041] [PMID: 34564303]
- [20] Srivastava D, Rajitha B, Agarwal S, Singh S. Pattern-based image retrieval using GLCM. *Neural Comput Appl* 2020; 32(15): 10819-32.
[http://dx.doi.org/10.1007/s00521-018-3611-1]
- [21] Limkin EJ, Reuzé S, Carré A, *et al.* The complexity of tumor shape, spiculatedness, correlates with tumor radiomic shape features. *Sci Rep* 2019; 9(1): 4329.
[http://dx.doi.org/10.1038/s41598-019-40437-5] [PMID: 30867443]
- [22] Ramli Z, Karim MKA, Effendy N, *et al.* Stability and reproducibility of radiomic features based on various segmentation techniques on cervical cancer DWI-MRI. *Diagnostics* 2022; 12(12): 3125.
[http://dx.doi.org/10.3390/diagnostics12123125] [PMID: 36553132]
- [23] Xu L, Wang MY, Qi L, Zou YF, Fei-Yun WU, Sun XL. Radiomics approach to distinguish between benign and malignant soft tissue tumors on magnetic resonance imaging. *Eur J Radiol Open* 2024; 12: 100555.
[http://dx.doi.org/10.1016/j.ejro.2024.100555] [PMID: 38544918]
- [24] Nahm FS. Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean J Anesthesiol* 2022; 75(1): 25-36.
[http://dx.doi.org/10.4097/kja.21209] [PMID: 35124947]
- [25] Conti A, Duggento A, Indovina I, Guerrisi M, Toschi N. Radiomics in breast cancer classification and prediction. *Semin Cancer Biol* 2021; 72: 238-50.
[http://dx.doi.org/10.1016/j.semcancer.2020.04.002] [PMID: 32371013]
- [26] Li H, Zhu M, Jian L, *et al.* Radiomic score as a potential imaging biomarker for predicting survival in patients with cervical cancer. *Front Oncol* 2021; 11: 706043.
[http://dx.doi.org/10.3389/fonc.2021.706043] [PMID: 34485139]
- [27] Chitalia RD, Kontos D. Role of texture analysis in breast MRI as a cancer biomarker: A review. *J Magn Reson Imaging* 2019; 49(4): 927-38.
[http://dx.doi.org/10.1002/jmri.26556] [PMID: 30390383]
- [28] Fusco R, Di Bernardo E, Piccirillo A, *et al.* Radiomic and artificial intelligence analysis with textural metrics extracted by contrast-enhanced mammography and dynamic contrast magnetic resonance imaging to detect breast malignant lesions. *Curr Oncol* 2022; 29(3): 1947-66.
[http://dx.doi.org/10.3390/currenol29030159] [PMID: 35323359]
- [29] Wang K, An Y, Zhou J, Long Y, Chen X. A novel Multi-Level feature selection method for radiomics. *Alex Eng J* 2023; 66: 993-9.
[http://dx.doi.org/10.1016/j.aej.2022.10.069]
- [30] Mao N, Yin P, Zhang H, *et al.* Mammography-based radiomics for predicting the risk of breast cancer recurrence: A multicenter study. *Br J Radiol* 2021; 94(1127): 20210348.
[http://dx.doi.org/10.1259/bjr.20210348] [PMID: 34520235]
- [31] Cui P, Athey S. Stable learning establishes some common ground between causal inference and machine learning. *Nat Mach Intell* 2022; 4(2): 110-5.
[http://dx.doi.org/10.1038/s42256-022-00445-z]
- [32] Uçar MK, Nour M, Sindi H, Polat K. The effect of training and testing process on machine learning in biomedical datasets. *Math Probl Eng* 2020; 2020: 1-17.
[http://dx.doi.org/10.1155/2020/2836236]
- [33] Polo TCF, Miot HA. Use of ROC curves in clinical and experimental studies. *J Vasc Bras* 2020; 19: e20200186.
[http://dx.doi.org/10.1590/1677-5449.200186] [PMID: 34211533]
- [34] Khan AA, Mahendran RK, Perumal K, Faheem M. Dual-3DM3AD: Mixed transformer based semantic segmentation and triplet pre-processing for early multi-class Alzheimer's diagnosis. *IEEE Trans Neural Syst Rehabil Eng* 2024; 32: 696-707.
[http://dx.doi.org/10.1109/TNSRE.2024.3357723] [PMID: 38261494]