KMANPUB Health Nexus Vol. 4 No. 3 (2026)

Article history:
Received 14 February 2026
Revised 15 June 2026
Accepted 23 June 2026
Published online 01 July 2026

Latency-Aware Breast Cancer Detection in Mammography Images Using a Fog-Cloud Framework Based on Stacked Transfer Learning and PSO-Optimized XGBoost

Zahraa Abdulmajeed Ibrahim Al-Mohammed[1], Esmaeil Bagheri[2*], Ameer Hussein Mohammed Ali[3], Mehdi Hamidkhani[4]

1. PhD Student, Department of Computer Engineering, Isf.C., Islamic Azad University, Isfahan, Iran
2. Department of Engineering, Deh.C., Islamic Azad University, Isfahan, Iran
3. Al-Najaf Technical Institute, Al-Furat Al-Awsat Technical University, Najaf, Iraq
4. Department of Electrical Engineering, Isf.C., Islamic Azad University, Isfahan, Iran

*** Corresponding author email address:** bagheri@iau.ac.ir

Editor	Reviewers
Shokouh Navabinejad Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. sh.navabinejad@kmanresce.ca	Reviewer 1: Mehdi Rostami Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. Email: dr.mrostami@kmanresce.ca Reviewer 2: Yaghob Badriazarin Associate Professor of Sport Sciences, Tabriz University, Tabriz, Iran. Email: badriazarin@tbzmed.ac.ir

1. Round 1

1.1 Reviewer 1

Reviewer:

In Section 2, the sentence “Mammography studies have shown that transfer-learning models, when combined with preprocessing and data augmentation, can be more stable than complete training from scratch” needs more precise operationalization. Please specify what “more stable” means in this context: higher cross-validation consistency, lower variance across folds, better external generalization, reduced overfitting, or improved calibration. Without defining stability, the statement remains broad and difficult to evaluate scientifically.

In Section 3.3, the paragraph on stacking correctly notes that “out-of-fold outputs were used to train the second and final levels.” However, the exact stacking protocol is not sufficiently reproducible. Please describe step-by-step how out-of-fold predictions were generated for each base model, how second-level classifiers were trained, whether nested cross-validation was used for PSO tuning, and how the final XGBoost meta-classifier was trained without seeing predictions derived from models trained on the same samples.

In Section 3.6, the objective function is given as "Score = αAccuracy + βPrecision + γRecall + δF1," with a balanced configuration of α = β = γ = δ = 0.25. This formulation is underspecified because precision, recall, and F1 may be macro-averaged, weighted-averaged, class-specific, or focused on the malignant class. Please define exactly which form of each metric was used in PSO fitness evaluation and justify whether malignant-class recall should receive higher weight given the clinical cost of false negatives.

In Section 4.2, the authors state that "The output of each base model was extracted as a two-class probability vector," producing a six-dimensional vector. This suggests that only final probabilities, rather than deep feature embeddings, were used for stacking. Please clarify whether the term "feature extraction" refers to extraction of internal CNN embeddings or simply classification probabilities. If only probabilities were used, the manuscript should avoid overstating "deep feature extraction" at the stacking stage or should add experiments using penultimate-layer feature vectors.

Authors revised the manuscript and uploaded the updated document.

1.2 Reviewer 2

Reviewer:

In Table 1, the row describing the "Present method" lists "Combination of accuracy and low-latency deployability" as the key point. This should be revised because the manuscript later acknowledges that the framework is latency-aware rather than fully validated as low-latency. Please align Table 1 with the more cautious wording used later in the manuscript and avoid claiming low-latency deployability unless end-to-end latency has been empirically benchmarked against cloud-only and fog-only baselines.

In Section 3.1, the manuscript itself notes that "The exact number of images retained after quality control, the benign/malignant distribution before and after exclusions, and whether classification was performed on full mammograms, cropped ROIs, or both should be explicitly reported." This is a major methodological gap that must be resolved before publication. Please provide the initial dataset size, exclusion criteria, final included sample size, class distribution, lesion type distribution if available, and whether the model used full mammograms, ROI crops, patches, or a combined input strategy.

In Section 3.1, the sentence "Following initial quality control, each sample entered the preprocessing pipeline and was then assigned to one of the cross-validation folds" raises a potential data leakage concern. If multiple images, views, or ROIs belong to the same patient or lesion, random sample-level splitting may allow related images from the same case to appear in both training and test folds. Please clarify whether splitting was performed at the patient/case level rather than the image level, and revise the protocol accordingly if necessary.

In Section 3.2, the paragraph states that "all images were resized to 224 × 224 pixels," including the branch using EfficientNetB7. This requires stronger justification because mammography contains subtle microcalcifications and fine lesion margins that may be degraded by aggressive downsampling. Please report whether alternative resolutions were tested, whether ROI-level preservation was used, and how the authors ensured that clinically relevant small structures were not lost during resizing.

In Section 3.2 and Algorithm 1, the preprocessing pipeline includes Gaussian filtering followed by median filtering, but the manuscript does not report filter kernel size, sigma value, order of operations, or whether these parameters were optimized.

Health Nexus
E-ISSN: 2981-2569

Because excessive denoising can blur lesion boundaries or microcalcifications, please provide the exact preprocessing parameters and justify them through either prior literature, ablation testing, or visual quality assessment.

In Section 3.3, the manuscript states that "The data were divided using stratified five-fold cross-validation," but later reports an "unseen test subset containing 102 benign and 96 malignant samples." The relationship between cross-validation and the final test subset is unclear. Please specify whether the 198 samples represent one held-out test split, the aggregated test predictions across five folds, or a separate test set after cross-validation-based model selection. This distinction is essential for interpreting the reported 97.5% accuracy.

Authors revised the manuscript and uploaded the updated document.

2. Revised

Editor's decision after revisions: Accepted.

Editor in Chief's decision: Accepted.

Health Nexus
E-ISSN: 2981-2569