Efficient Breast Segmentation and Pectoral Muscle Removal Technique in Wavelet Domain for Computer-Aided Diagnosis

Muhammad Nadeem
Department of Computer Science & Information Technology, University of South Punjab (USP), Multan, Punjab, Pakistan.
Email: mahrnadeem658@gmail.com

Babar Bakht Khan
Department of Computer Science & Information Technology, University of South Punjab (USP), Multan, Punjab, Pakistan.
Email: babarbakht@gmail.com

Ghulam Muhy Ud Deen Raee
Department of Computer Science & Information Technology, University of South Punjab (USP), Multan, Punjab, Pakistan.
Email: gmraees@gmail.com

Muhammad Shahid
Department of Computer Science & Information Technology, University of South Punjab (USP), Multan, Punjab, Pakistan.
Email: mshahidbhatti0007@gmail.com

Muhammad Talah Zubair
Department of Computer Science, NCBA&E Lahore Sub Campus Multan, Punjab, Pakistan.
Email: talhazubair009@gmail.com

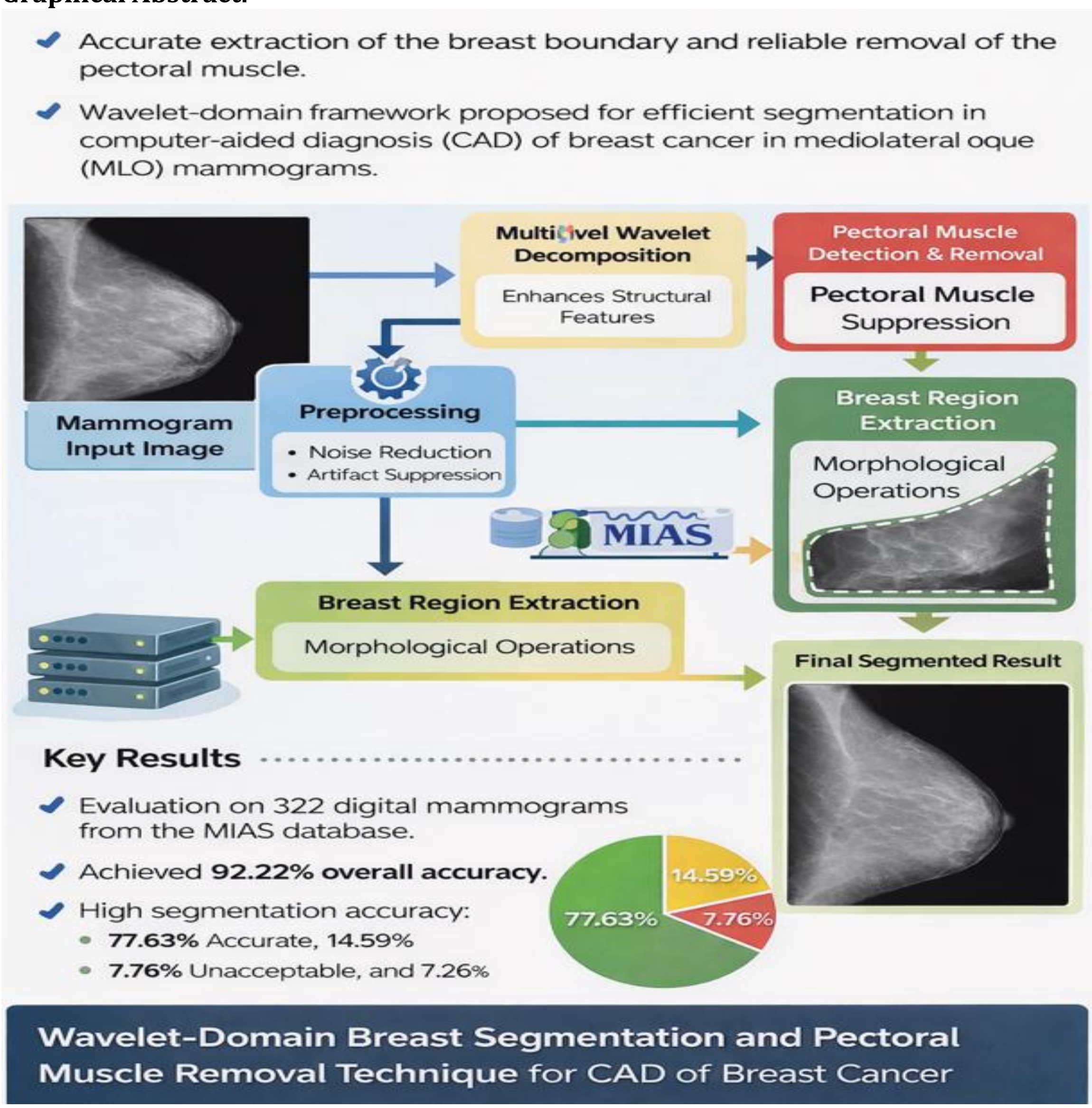
Received on: 10-10-2025 Accepted on: 15-11-2025

Abstract

Accurate extraction of the breast boundary and reliable removal of the pectoral muscle are essential preprocessing steps for computer-aided diagnosis of breast cancer using mammographic images. In this work, a wavelet-domain framework is proposed to effectively segment the breast region and identify the pectoral muscle in mediolateral oblique (MLO) mammograms. The breast region is first delineated using multilevel wavelet decomposition, which enhances structural features while suppressing background noise and imaging artifacts. Subsequently, morphological

operations are employed to refine the segmented region and to detect and eliminate the pectoral muscle. The proposed method was evaluated using 322 digital mammograms from the MIAS database, demonstrating high accuracy and robust performance for breast region extraction and pectoral muscle removal.
Keywords: Pectoral muscle, Mediolateral oblique, Cranial-Caudal, Computer-aided detection system

Graphical Abstract:



1: Introduction

Breast cancer remains one of the leading causes of cancer-related mortality among women worldwide, with the number of fatalities continuing to rise due to late diagnosis and metastasis to distant organs [1]. Early detection plays a decisive role in reducing mortality and improving treatment outcomes. In this context, Computer-Aided Detection (CAD) systems have emerged as powerful tools for assisting radiologists in the early identification of suspicious lesions in mammographic images, thereby enhancing diagnostic accuracy and reducing human error [2]. A number of techniques have been proposed to improve the reliability of preprocessing steps in mammogram analysis, particularly for the removal of the pectoral muscle region. A novel method for pectoral muscle suppression was introduced in [3,4], while the use of cartoon–texture decomposition has been explored for accurate localization of the muscle boundary [5]. Breast boundary detection based on homogeneous region contours has also been investigated in [6,7]. Furthermore, the integration of morphological operators with polynomial modeling has been reported as an effective strategy for pectoral muscle detection [8]. Several additional approaches have been proposed in the literature for pectoral muscle removal and breast region segmentation [9,10]. For example, graph theory combined with active contour models has been employed to remove the pectoral muscle [11], and Bézier curve fitting has been utilized to smooth the muscle boundary by tracking edge points [12]. A normalized graph-cut framework was designed to isolate the pectoral muscle edge [13], while shortest-path based contour detection has also shown promising performance [14]. In addition, bit-depth reduction combined with wavelet transformation has been applied for pectoral muscle suppression in mammograms [15]. In mediolateral oblique (MLO) mammographic views, the pectoral muscle typically appears as a high-intensity triangular region occupying a significant portion of the image. Although it contains anatomical information, its presence adversely affects breast segmentation, feature extraction, and classification processes, often leading to an increased false-positive rate. Accurate removal of this region is therefore a crucial preprocessing step. However, the task remains challenging due to variations in muscle shape, size, intensity distribution, and its similarity to surrounding breast tissue.

To address these challenges, this study proposes a robust framework for breast border extraction and pectoral muscle removal using mathematical morphology in the wavelet domain. The remainder of this paper is organized as follows: Section 2 presents the proposed methodology along with its graphical representation, Section 6 discusses the performance evaluation, and Section 7 concludes the study.

2: Proposed Methodology

Accurate identification and removal of the pectoral muscle remains a challenging task in mammographic image preprocessing. Many existing techniques still suffer from limited robustness and reduced accuracy, particularly in complex cases where the intensity of the pectoral muscle is similar to dense breast tissue. These limitations can significantly mislead downstream tasks such as tumor segmentation, feature extraction, and classification, ultimately reducing the reliability of computer-aided diagnosis systems. Consequently, there is a strong need to develop an improved and more reliable framework that enhances segmentation precision and minimizes false detections. To address these challenges, this study introduces a robust wavelet-domain based framework for breast border extraction and pectoral muscle removal. The proposed pipeline integrates multilevel wavelet decomposition, artifact suppression, and morphology-based refinement to achieve accurate segmentation. The method is designed as a sequential and automated workflow, ensuring consistent performance across mammographic images with varying contrast, noise levels, and anatomical variations. The overall workflow of the proposed system is illustrated in Figure 1.

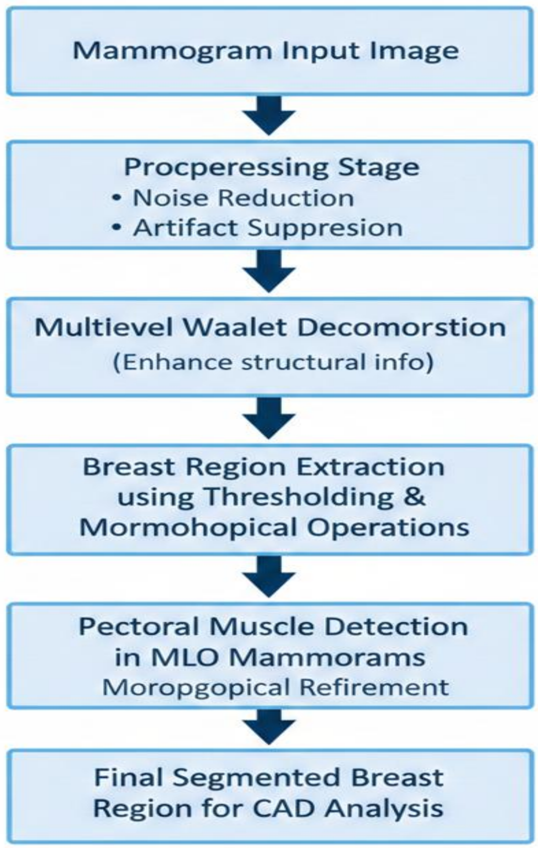


Figure 1: Sequential workflow of the proposed wavelet-domain framework for breast border extraction and pectoral muscle removal in mammographic images.

3: Removal of pectoral muscle

Mammographic images typically consist of two main regions: the breast tissue region and the non-breast background region. The background often contains radiopaque artifacts such as labels, scanning markers, wedges, and acquisition annotations that do not contribute to diagnostic interpretation. These high-intensity objects can negatively influence segmentation accuracy if not removed during preprocessing. To address this issue, a seed-point based strategy is employed to identify and suppress high-intensity background artifacts while preserving the true breast tissue region. The effectiveness of this preprocessing step is illustrated in Figure 2, which presents mammograms before and after artifact removal.

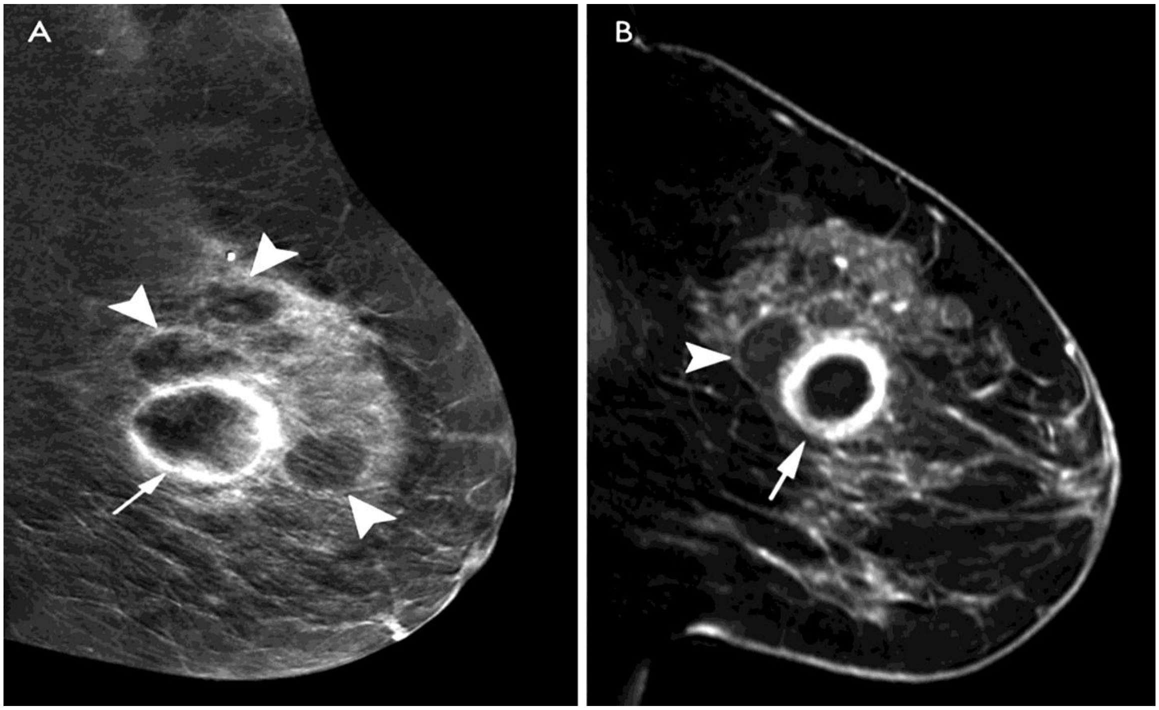


Figure 2. (a) Mammograms before and (b) after.

In addition to artifact suppression, mammographic images are commonly affected by quantum noise introduced during the image acquisition process. Many heuristic or conventional denoising filters tend to oversmooth the image and degrade important anatomical edges, which are critical for accurate boundary detection [16]. To preserve structural details while effectively reducing noise, a median filtering approach is adopted in this work. The median filter is widely recognized for its ability to remove impulsive and speckle noise without significantly blurring edges. The mathematical formulation of the median filtering operation used in this study is presented in Equation (1).

$\nu\, \iota . \kappa \;=\; arg \sum \sum_{\gamma \cdot}^{Z} = 1 \| \omega m \;-\; \omega \gamma \| \; \rho,$ ....(1)

Where υικ is the output value for location χικ, νικ=ν1 , ν2 , ν3 , ...νZ, ι,κ are representing the position in a 2D image, m and γ are filtering window indices, where filtering window length is denoted by Z, and ρ denote the Lρ norm. Commonly, ρ=1 corresponds to standard median filtering. The main purpose of this preprocessing is to obtain significant mammograms that further help us to obtain accurate breast profile. We proposed a unique method to for removal of pectoral muscle in order to achieve accurate tumor region.

4: Enhancement of salient feature of pectoral muscle

A mammographic image generally consists of two primary regions: the breast tissue region and the surrounding background. The background often includes radiopaque artifacts such as labels, markers, wedges, and other acquisition-related annotations that are irrelevant for diagnosis and may interfere with accurate segmentation. To eliminate these unwanted elements, a seed-point based approach is employed to enhance and isolate high-intensity background components, enabling their effective removal while preserving the breast tissue. The effectiveness of this preprocessing step is illustrated in Figure 2, which shows mammograms before and after artifact suppression. Mammographic images are also commonly affected by quantum noise introduced during the imaging process. Many heuristic denoising techniques tend to blur fine anatomical details and degrade important edge information [16]. To overcome this limitation, a median filtering technique is adopted due to its ability to suppress noise while preserving structural boundaries. The mathematical formulation of the median filter applied in this study is provided in Equation (1).

$$f = \sum_i \alpha i \varphi i \rightarrow \quad (2)$$

$$\varphi i = \left\{ \begin{matrix} 1 & 0 < \tau < \frac{1}{2} \\ 1 & \frac{1}{2} \leq \tau < 1 \\ 0 \; otherwise & \end{matrix} \right\} \ldots\ldots(3)$$

5: Extraction of breast border line

To accurately refine the breast parenchyma and suppress outlier contours, a contour-based refinement strategy combined with morphological operations is applied to the approximation sub-band obtained from wavelet decomposition. The approximation component primarily contains the low-frequency structural information of the mammogram, making it suitable for extracting the overall breast boundary. In contrast, the high-frequency detail sub-bands mainly capture edges, noise, and fine structures. To effectively suppress noise while preserving diagnostically relevant edges, the Minimum Mean Squared Error (MMSE) estimator is applied to the detailed sub-bands [19]. After denoising and refinement of both approximation and detail components, inverse wavelet reconstruction is performed using all four sub-bands to obtain an enhanced mammographic image. In the reconstructed image, high-intensity regions are amplified to facilitate accurate segmentation. A seed-based region growing approach is then employed to segment the breast region. The seed point is positioned within the convex hull of the breast area and iteratively eroded until it reaches the true anatomical boundary. This procedure ensures that the final segmented region conforms closely to the natural geometry of the breast while preventing leakage into the background. The proposed framework employs a two-level wavelet decomposition, commonly referred to as a dyadic scale representation. This multiresolution analysis enables efficient separation of structural and textural information, thereby improving segmentation accuracy and computational efficiency [2,3]. In general, the wavelet analysis of a signal β produces a set of coefficients ρ(ω,υ), which represent the signal in terms of scale and spatial location, as expressed in the following equation.

$$P(\omega, \gamma) = \sum_{\in \partial \mu} \beta \mu \sigma \; \omega, \gamma_{\mu \rightarrow} \quad \ldots\ldots(4)$$

6: Experiments and Results

Let a mammographic image be represented as a set of pixels ω={ω1,ω2,...,ωn}, where the total number of pixels |ω||\omega||ω| corresponds to the product of the number of rows and columns of the image matrix (row × col). This representation enables a pixel-wise analysis of the breast region and facilitates the implementation of segmentation and evaluation procedures within a discrete image domain. The performance of the proposed framework was evaluated using a total of 322 mammographic images (μ) obtained from the MIAS (Mammographic Image Analysis Society) database. This dataset contains mediolateral oblique (MLO) mammograms with varying tissue densities, contrast levels, and anatomical variations, making it suitable for validating the robustness and generalization capability of the proposed method. For performance assessment, a subjective evaluation strategy was adopted in addition to quantitative analysis. In the subjective assessment, the segmentation results were visually inspected and validated by an experienced radiologist to ensure clinical relevance and accuracy. The radiologist examined the correctness of the extracted breast boundary and the effectiveness of pectoral muscle removal. The summarized evaluation outcomes are presented in Table 1.

Table 1. Evaluation of the proposed segmentation technique

Category	Evaluation Criteria	No. of Images	Percentage
Accurate (α)	$I^{v}_{sg} \equiv I^{k}_{g\tau}$	250	77.63%
Acceptable (ρ)	$I^{v}_{sg} > \frac{1}{2}(I^{k}_{g\tau})$	47	14.59%
Unacceptable (γ)	$I^{v}_{sg} \leq \frac{1}{2}(I^{k}_{g\tau})$	25	7.76%
Total images (μ)	—	**322**	—
Overall Accuracy (Aσ)	$A\omega = \frac{\sum_{i=1}^{\mu}(\alpha + \rho)}{\mu}$	—	**92.22%**

7: Conclusion

In this study, a reliable and consistent approach for the segmentation of the pectoral muscle was developed and evaluated using the widely recognized MIAS mammographic image database. For accurate analysis and validation, the true pectoral muscle boundaries were manually delineated by an experienced radiologist. These expert-annotated contours were used as the ground truth reference for assessing the performance of the proposed automated framework. The comparison between the automatically detected boundaries and the manually annotated ground truth demonstrates that the proposed method achieves accurate and consistent pectoral muscle segmentation across mammograms with varying tissue densities and anatomical variations. The experimental findings indicate that the proposed approach provides promising and clinically relevant results, making it a valuable preprocessing step for computer-aided diagnosis systems. By improving the accuracy of breast region extraction and reducing false positives, the method shows strong potential for supporting early-stage breast cancer detection.

Competing Interests

There are no conflicts of interest exist among authors of this research paper.

ACKNOWLEDGEMENTS

Author Contributions

M.N, B.B.K, M.S, G.M.U.D.R and M.T.Z wrote the main manuscript Draft, writing, studied Data validation and Editing, Reviewing. All authors reviewed the manuscript outline.

Funding

No external fundings.

Data Availability

No data sets were generated or analysed during the current study.

References:

[1] Abbas, M. A., Mahar, J., Hameed, N., & Rasool, M. S. (2025). DFT-Guided Design of a Low-Band-Gap Pyrazoline Scaffold: The Critical Role of a Para-Nitro Substituent. *Multidisciplinary Surgical Research Annals*, *3*(3), 461-503.

[2] Abbas, M. A., Mahar, J., Khan, M. J., Rasool, M. S., & Khan, M. Z. (2025). IN SILICO INVESTIGATION OF 3, 6-DIPHENYL-[1, 2, 4] TRIAZOLO [3, 4-B][1, 3, 4] THIADIAZOLE DERIVATIVES AS EGFR MODULATORS FOR LUNG CANCER TREATMENT. *The Cancer Research Review*, *4*(2), 243-308.

[3] Abbas, M. A., Mahar, J., Rasool, M. S., Khan, M. J., & Khan, M. Z. (2025). The Dual Therapeutic Promise of Quinoa: Exploring Antidiabetic and Antioxidant Effects through Experimental and Computational Models. *Multidisciplinary Surgical Research Annals*, *3*(3), 504-544.

[4] Abbas, M. A., Junaid, M. J. M., Rasool, M. S., & Mahar, J. (2025). Structural and NLO Properties of Novel Organic 4-Bromo-4-Nitrostilbene Crystal: Experimental and DFT Study. *International Research Journal of Management and Social Sciences*, *6*(4), 1-20.

[5] Abbas, M. A., Khan, M. Z., Atif, H. M., Shahzad, A., & Mahar, J. (2025). Computer-Aided Analysis of Oxino-bis-Pyrazolederivative as a Potential Breast Cancer Drug Based on DFT, Molecular Docking, and Pharmacokinetic Studies: Compared with the Standard Drug Tamoxifen. *Indus Journal of Bioscience Research*, *3*(6), 535-537.

[6] Ghafoor, A., Abbas, M. A., Ahmad, J., Ali, R., Abbas, M., Ahmad, A., ... & Masood, H. T. (2025). Synthesis and characterization of pyranopyrazole derivatives as potential organic nonlinear optical (NLO) materials and density functional theory (DFT) studies. *FRONTIERS IN CHEMICAL SCIENCES*, *6*(2), 103-116.

[7] Rasool, M. S., Abbas, M. A., Khan, M. J., Mahar, J., & Khan, M. Z. IDENTIFICATION OF NATURAL EGFR TYROSINE KINASE INHIBITORS FROM CHENOPODIUM QUINOA WILLD. VIA COMBINATORIAL IN SILICO AND PHARMACOLOGICAL SCREENING.

[8] Altaf, R., S. Altaf, M. Hussain, R.U. Shah, R. Ullah, M.I. Ullah, and R. Datta. (2021). Heavy metal accumulation by roadside vegetation and implications for pollution control. PLoS ONE, 16(5): e0249147.

[9] Chen, H., R. Yang, X. Zhang, Y. Chen, Y. Xia and X. Xu. (2021). Foliar application of gibberellin inhibits the cadmium uptake and xylem transport in lettuce (Lactuca sativa L.). Scientia Horticulturae, 288: 110410.

[10] Hussaini, K.M., J. Akhtar, M. Anwar-ul-Haq and A.A. Maann. (2022). Impact of heavy metals-enriched dust particles on the growth, metal accumulation and gas exchange attributes of various plant species in urban areas of Faisalabad, Pakistan. Pakistan Journal of Agricultural Sciences, 59(6): 1011–1020.

[11] Korzeniowska, J. and E. Panek. (2019). Trace metal concentrations in Pleurozium schreberi and Taraxacum officinale along road No. 7. Ecological Chemistry and Engineering, 26(4): 651–663.
[12] Kumar, V., J. Singh and P. Kumar. (2019). Heavy metals accumulation in crop plants: Sources, response mechanisms, stress tolerance and their effects. Contaminants in Agriculture and Environment: Health Risks and Remedies, 1: 38–57.
[13] Kumar, A., P. Kumar, H. Singh and N. Kumar. (2021). Adaptation and mitigation potential of roadside trees with bio-extraction of heavy metals under vehicular emissions and their impact on physiological traits during seasonal regimes. Urban Forestry & Urban Greening, 58: 126900.
[14] Kumar, A., H. Singh, G. Kumari, S. Bisht, A. Malik, N. Kumar, and R. Kaushal. (2022). Adaptive resilience of roadside trees to vehicular emissions via leaf enzymatic, physiological, and anatomical trait modulations. Environmental Pollution, 313: 120191.
[15] Luo, S., A. Calderón-Urrea, J. Yu, W. Liao, J. Xie, J. Lv, and Z. Tang. (2020). The role of hydrogen sulfide in plant alleviates heavy metal stress. Plant and Soil, 449: 1– 10.
[16] Martini, A.N., M. Papafotiou, I. Massas and N. Chorianopoulou. (2022). Using the halophyte Crithmum maritimum in green roofs for sustainable urban horticulture: Effect of substrate and nutrient content analysis including potentially toxic elements. Sustainability, 14(8): 4713.
[17] Ray, S. and M. Neogy. (2022). Identification of vehicular emission hyper accumulator and its subsequent use as phytoremediation. Bangladesh Journal of Botany, 51(2): 207–216.
[18] Singh, H., S. Bandyopadhyay, A. Kumar, M. Singh, A. Malik, G. Kumari, and P.S. Rawat. (2022). Understanding the physiological and biophysical response of urban roadside plantations for assessing adaptation and mitigation mechanisms toward vehicular emissions. Urban Climate, 44: 101183.
[19] Yoon, J., X. Cao, Q. Zhou and L.Q. Ma. (2006). Accumulation of Pb, Cu, and Zn in native plants growing on a contaminated Florida site. Science of the Total Environment, 368: 456–464.
[20] Yu, G., J. Ma, P. Jiang, J. Li, J. Gao, S. Qiao and Z. Zhao. (2019). IOP Conference Series: Earth and Environmental Science, 310: 052004