Indonesian Journal of Electrical Engineering and Computer Science
Vol. 36, No. 3, December 2024, pp. 1613~1619
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v36.i3.pp1613-1619

Developed improved lion optimization for breast cancer classification using histopathology images

Pattan M. D. Ali Khan, Xavier Arputha Rathina
Department of CSE, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, India

Article Info

Article history:

Received May 15, 2024
Revised Aug 19, 2024
Accepted Aug 26, 2024

Keywords:

Breast cancer
Convolutional neural network
Deep learning
Histopathological database
Improved lion optimization
Uniform experimental design

ABSTRACT

Breast cancer, a prevalent kind of cancer, is a major health problem among women. Researchers recently achieved categorization effectiveness of breast cancer (BC) detection in histopathology picture database using convolutional neural networks (CNNs) of medical image processing. Although CNN method parameter settings were complex, employing breast cancer histopathological database (BCHD) data for categorization was valued as expensive. This research used uniform experimental design (UED) to solve these issues and improved lion optimization (ILO) breast cancer histopathology image categorization. To optimize the variables at UED-ILO, a regression method was employed. According to the experimental data, the proposed approach of UED-ILO (uniform experimental design based improved lion optimization) variable optimization provided a categorization accuracy rate of 84.41%. Finally, the proposed approach can effectively increase classification accuracy, with results that outperform others of an equivalent nature.

Corresponding Author:

Pattan M. D. Ali Khan
Department of CSE, B.S Abdur Rahman Crescent Institute of Science and Technology
Chennai, India
Email: alik_12@rediffmail.com

1. INTRODUCTION

Breast cancer (BC) is the most frequent malignancy in women worldwide. In Taiwan (population 23 million), 1 in 120 women was recognized with BC each year, and the prevalence of BC was rising [1]. Histopathological image categorization accuracy is critical for early BC detection. BC diagnosis relies on histopathological imaging methods such as ultrasound, positron emission tomography (PET), surgical incision mammography, thermography, and magnetic resonance imaging (MRI) [2], [3]. Diagnosing BC using histopathological pictures was challenging in the early stages since these pictures can transmit warning signals and symptoms [4]. As a result, a variety of computer-assisted solutions was created to address the shortcomings of histological picture evaluation [5]. The research team was the visual geometry group (VGG) or Google of recently applied advanced engineering methods to build the VGG-16, residual network (ResNet), or general nursing and midwifery (GNM) [6]. Deep learning (DL) models based on convolutional neural networks (CNN) are among the technical methods utilized to enhance the efficiency of breast cancer prediction [7]-[9]. This was accomplished by hypervariable optimization. To increase the precision attained to previous research employing BreakHis of picture categorization, this research employs DL networks as a CNN with variable optimization. DL networks generate amazing outcomes to picture evaluation applications in medicine [10].

The LeNet-5 system is a powerful network of CNN applications that has a high recognition rate. LeNet-5 (RMSprop-root mean squared propagation) [11], LeNet-5 (SGDM-stochastic gradient descent with momentum) [12], LeNet-5 (Adam-analysis data model) [13], and single-layer CNN [14], and random forest (RF) classifier+PFTAS [15] are some other DL networks. To solve practical issues, CNN training necessitates extensive calculation with high sample sizes and variable settings [16]. This reduces the number of networks calculating samples while meeting the essential parameter values for CNN applications. Experimentation is used to identify the most appropriate CNN model [17], [18]. Latchoumi *et al.* [19] also employed a uniform experimental design (UED) to determine the network's variables which enhanced evaluation. The CNN variables used for BC histopathology image categorization, on the other hand, are optimized using a UED in this study [20].

UED an approach that was statistical experimental design mathematical statistics, and probability theory, [21] was utilized to optimize CNN variables and reduce experiment computation time. UED could be utilized to identify appropriate sample groups and distribute all available experimental variables obtained in some studies tests [22]. According to a collection of UED tables, the number of layers equals the number of test runs [23]. Furthermore, to obtain appropriately precise projections, UED examines the variables impacting the findings in a restricted number of experiments [24].

The UED approach is employed in this study to optimize the CNN approach parameters to the application of BC histopathology picture categorization [25]. The current work created a CNN depending on UED to overcome the difficult variable setup issue to improve classification accuracy. As an outcome, the key contribution of this research was to apply the UED technique to find the best CNN architectural variable combination for doing the fewest number of tests and taking the least amount of time. The regression evaluation was employed in the UED approach to determine optimization variables in CNN architecture. The proposed CNN-based uniform experimental design-based improved lion optimization (UED-ILO) variable optimization method surpasses current methods with high accuracy, making it more useful in clinical diagnostics.

2. PROPOSED SYSTEM

The proposed technique combines a CNN with unsupervised evolutionary DL using indirect encoding with local optimization (UED-ILO). This fusion aims to enhance the classification performance of the model by leveraging the strengths of both CNNs and evolutionary algorithms. Figure 1 illustrates the architectural layout of the proposed technique, showcasing how the CNN and UED-ILO components are integrated to achieve improved classification accuracy.

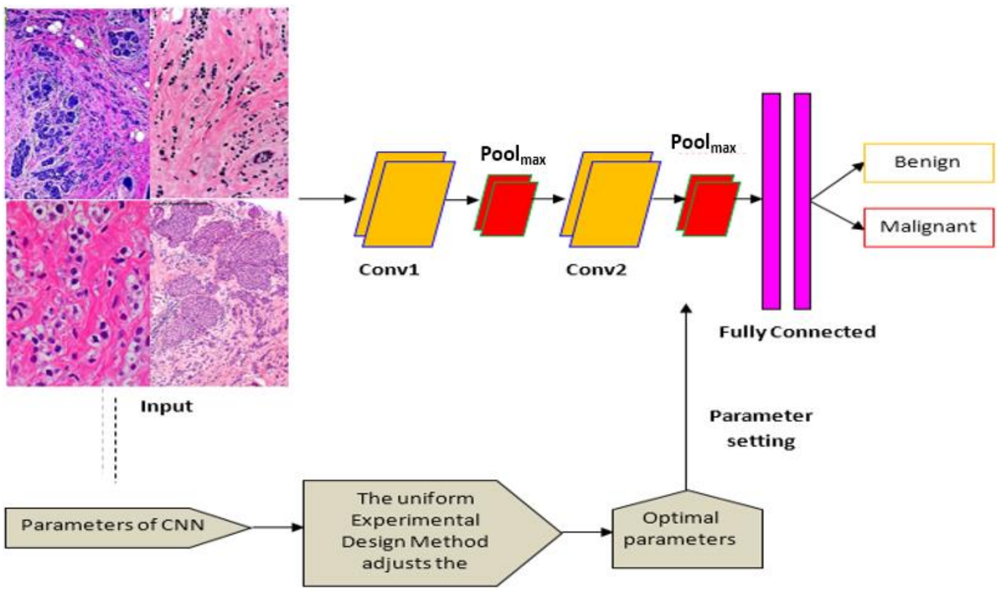
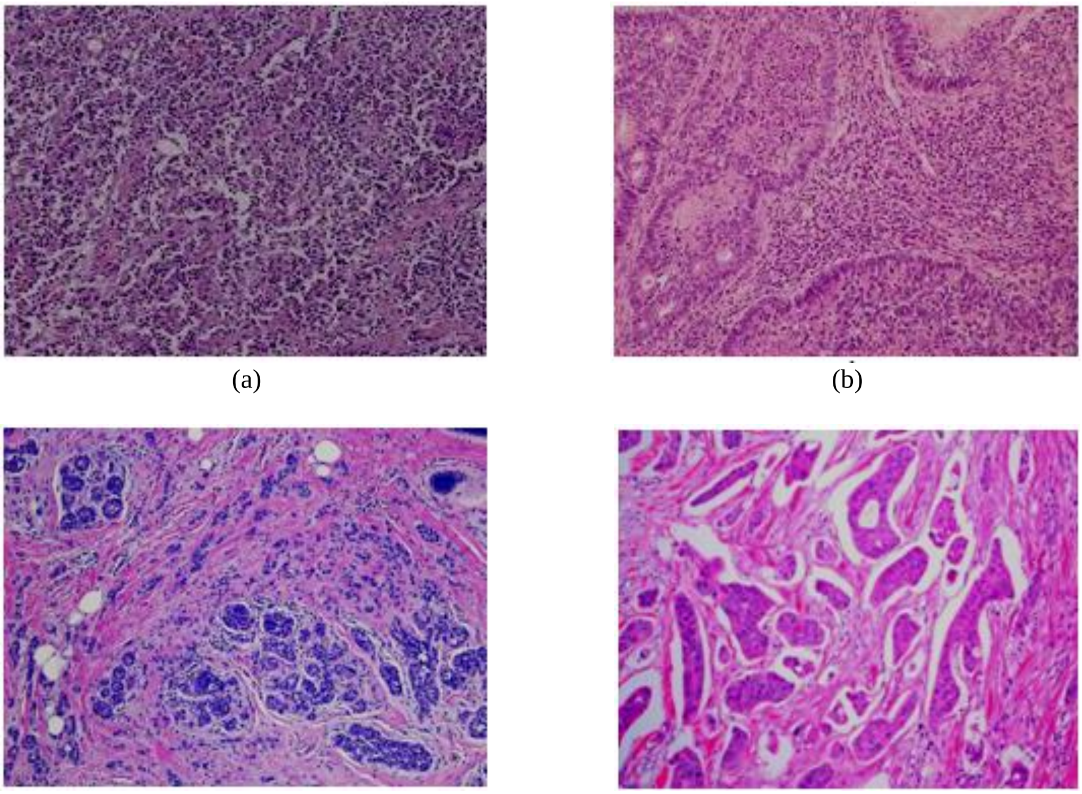


Figure 1. The proposed approach for histopathological image categorization of breast cancer

2.1. Materials

The experimental pictures in this investigation were generated through breast tissue samples and BreakHis were gathered from 82 people in the Brazilian Cytopathology Laboratory and pathological anatomy. Multiple breast tissue samples were aspirated in the operating room for each patient using a tiny biopsy needle. Each sample was created in the following manner: to retain the original tissue structure and molecular makeup, implantation in paraffin blocks and formalin fixation were conducted first. The paraffin blocks were then sliced into the 3-µm-thick sections using a high-power microwave (HPM). Lastly, the divisions were displayed on coated glass slides and viewed using a light microscope. BreakHis includes 7,909 700×460-pixel histopathological photos of BC at four rising magnifications (40×, 100×, 200×, and 400×), with 5,429 and 2,480 pictures of cancerous and benign tumors, correspondingly shown in Figure 2, further Figure 2(a) shows the magnification at 40×, Figure 2(b) shows the magnification at 100x, Figure 2(c) shows the magnification at 200x, Figure 2(d) shows the magnification of 400×. Although the class imbalance issue could skew CNN classification's discriminative capability; this is the BreakHis database restriction, and it would incline to predict pictures as malignant as shown in Table 1. As a result, the gathered data were separated into validation and training groups, with the first 70% of photographs used to train the network and the balance of 30% used to validate.



(a) (b)

(c) (d)

Figure 2. Histopathological pictures of BC: (a) 40×, (b) 100×, (c) 200×, and (d) 400× magnification

Table 1. BreakHis database picture segregation by class

Magnification	Malignant	Benign	Total
40x	1,375	654	2,002
100x	1,444	646	2,085
200x	1,402	628	2,015
400x	1,235	572	1,823
Total images	5,456	2,500	7,925

2.2. Algorithms for UED-ILO

This method is used to determine the best CC. The bubble-net serving method was used to hunt the animal. This approach consists of two basic phases: surrounding the prey and enveloping it with a bubble network. Searching for humpback lion food was done to reach a worldwide maximum/minimum. The procedure of nourishing a humpback lion with a bubble net is described in (1) and (2).

$$P_{m+1} = \begin{cases} P * m - EB & q < 0.5 \\ b^d . \cos(2\pi m) . E^1 + P * m & q \geq 0.5 \end{cases} \quad (1)$$

$$B^1 = [EP * m - pm] \quad (2)$$

Where P*-update the population-based best solution using an objective function. E and B are vector coefficient vectors and using (3) and (4). rd denotes the range vector range from 0 to 1. The weighted quantum fitness function can be calculated by using (5). Where M_{matrix} denotes membership matrix search agent, C_c signifies CC and the data elements clustered using the proposed system provide better results and enhance the fitness value. For lion optimization (LO), the logistics mapping (LM) method is used and determined by using (6). Algorithm 1 shows the UED-ILO algorithm.

$$E = 2.e.rd - e \quad (3)$$

$$B = 2.rd \quad (4)$$

$$F(P_x) = \frac{1}{Y(M_{matrix}.C_c)+1} \quad (5)$$

$$q_{k+1} = \partial q_k (1 - q_k) \quad (6)$$

Algorithm 1. Algorithm for UED-ILO

```
Input: Scanned images (MRI)
Output: Predict the histopathology
Begin
Step 1  Transformed into the greyscale image from the MRI image
Step 2  Segment the greyscale image using FCM
Step 3  Extract the features from the segmented image
Step 4  Steps to follow UED features are for classifying histopathology
            (i) Calculate the input features
```
$$FE(b(n)) = \sum_{q=1}^{k} wei_q \delta_q(n) \quad (7)$$
```
            (ii) Weighed Quantum estimation using (8)
```
$$wei = (l^T l)^{-1} l^T y \quad (8)$$
```
            (iii) RBF estimation
```
$$\delta_q(n) = exp\left[\frac{-|b(n) - cen_q|^2}{2\omega_q^2}\right] \quad (9)$$
```
Step 5  ILO algorithm to calculate the optimum value using (9)
End
```

Our algorithm proved susceptible to human biases and mistakes since it was trained on data that had previously been labeled by skilled pathologists. In addition to any anomalies or inconsistencies in our dataset, the scale of our dataset poses a serious issue for DL models. Large amounts of data have been used extensively in the learning of several large-scale models.

3. EXPERIMENTAL RESULTS

The three experiments may have used the same variable configurations, yet each observation was documented separately. To run 2 of this study, CNN's average classification accuracy is 83.4%. Furthermore, by utilizing UED-ILO for regression evaluation the average categorization accuracy of the optimized structure in BreakHis rose to 1.01% after CNN architecture factor optimization as shown in Table 2.

Table 2. The outcomes of a comparison of run 2 and the UED-ILO

Run	Parameter of factors								Performance in percentage
2	6	12	2	3	1	33	1	2	84.2
UED-ILO	7	12	1	3	1	9	2	2	85.13

The evaluation of the proposed UED-ILO factor optimization method utilizes a confusion matrix, depicted in Figure 3, to assess its classification performance. Within the 2,373 verification pictures, the confusion matrix reveals 1,997 accurate classifications, distinguishing between 521 benign and 1,476 cancerous cases. Additionally, the receiver operating characteristic (ROC) curve, illustrated in Table 3, showcases an area under the curve (AUC) value of 0.842, indicating the model's strong discriminative ability between benign and cancerous samples.

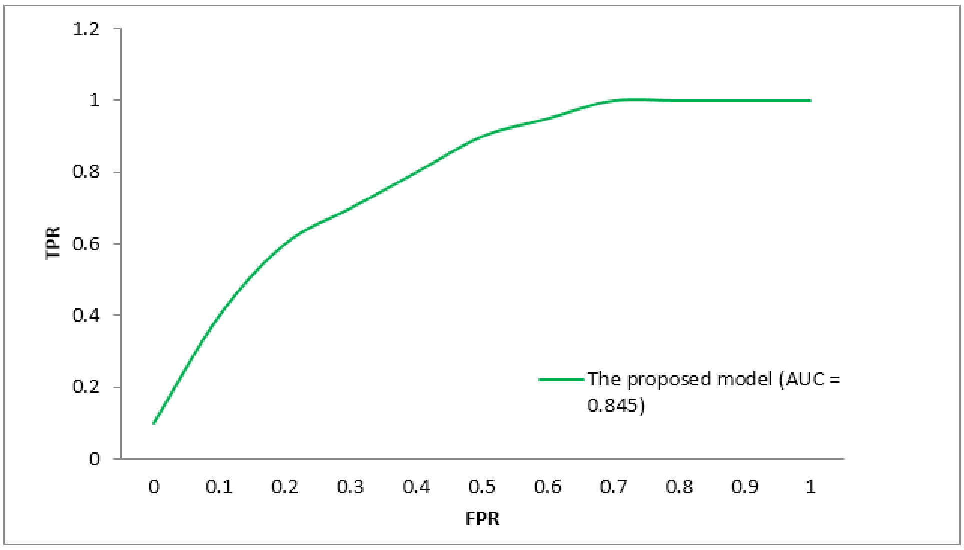


Figure 3. The proposed model's confusion and the proposed model's ROC curve

Table 3. Comparison results of the various methods

Method	Accuracy (%)
Multi-layer	77.55
PFTAS with RF classifier	82.33
SGDM with LeNet-5	80.72
ADAM with LeNet-5	82.25
RMS Prop with LeNet-5	82.61
CNN	83.20
CNN with UED-ILO	84.44

4. CONCLUSION

This article proposes optimizing UED-ILO with CNN parameters for deep learning networks utilized in studies on breast cancer histopathology image categorization. To optimize the variable combination, the recommended UED method used a regression evaluation and uniform experiment table to change the CNN framework. Testing findings showed that the proposed method's average predicted (with BreakHis) was 84.41%, which is 1.01% greater than the precision of a CNN employing UED-ILO. According to the experimental outcomes, the proposed UED-ILO variable optimization increases network efficiency and outperforms existing approaches. The benefits of the study include providing modeling users with a limited number of tests to establish optimal variable combinations of CNN design, reducing experimental timescales, and improving classification precision. The research's drawbacks include the inclusion of only the first and second convolutional levels as contributing variables. Nonetheless, UED-ILO variable optimization shows future learning capability and could analyze training and test data sets of varied sizes. Future research is concentrated on the ideal size of data regions in the generation of new architectural structures using deep learning algorithms. Furthermore, the disparity between malignant and benign datasets represents a weakness of the BreakHis dataset. As a result, in future studies, they will use the (GAN) method to extend the benign database.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who have supported and contributed to this research project. Primarily, we extend our heartfelt thanks to our guide for her unwavering guidance, invaluable insights, and encouragement throughout the research process.

REFERENCES

[1] M. M. Srikantamurthy, V. P. S. Rallabandi, D. B. Dudekula, S. Natarajan, and J. Park, "Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning," *BMC Medical Imaging*, vol. 23, no. 1, p. 19, Jan. 2023, doi: 10.1186/s12880-023-00964-0.

[2] D. Yu, J. Lin, T. Cao, Y. Chen, M. Li, and X. Zhang, "SECS: an effective CNN joint construction strategy for breast cancer histopathological image classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 2, pp. 810–820, Feb. 2023, doi: 10.1016/j.jksuci.2023.01.017.

[3] S. Majumdar, P. Pramanik, and R. Sarkar, "Gamma function based ensemble of CNN models for breast cancer detection in histopathology images," *Expert Systems with Applications*, vol. 213, p. 119022, Mar. 2023, doi: 10.1016/j.eswa.2022.119022.

[4] T. Araújo *et al.*, "Classification of breast cancer histology images using convolutional neural networks," *PLOS ONE*, vol. 12, no. 6, p. e0177544, Jun. 2017, doi: 10.1371/journal.pone.0177544.

[5] D. Banumathy, O. Ibrahim Khalaf, C. A. T. Romero, P. V. Raja, and D. K. Sharma, "Breast calcifications and histopathological analysis on tumour detection by CNN," *Computer Systems Science and Engineering*, vol. 44, no. 1, pp. 595–612, 2023, doi: 10.32604/csse.2023.025611.

[6] P. Agarwal, A. Yadav, and P. Mathur, "Breast Cancer prediction on breakhis dataset using deep CNN and transfer learning model," in *Lecture Notes in Networks and Systems*, vol. 238, 2022, pp. 77–88.

[7] D. P. Bhausaheb and K. L. Kashyap, "Detection and classification of breast cancer availing deep canid optimization based deep CNN," *Multimedia Tools and Applications*, vol. 82, no. 12, pp. 18019–18037, May 2023, doi: 10.1007/s11042-022-14268-y.

[8] J. Huang *et al.*, "BM-Net: CNN-based MobileNet-V3 and bilinear structure for breast cancer detection in whole slide images," *Bioengineering*, vol. 9, no. 6, p. 261, Jun. 2022, doi: 10.3390/bioengineering9060261.

[9] T. P. Latchoumi and L. Parthiban, "Quasi oppositional dragonfly algorithm for load balancing in cloud computing environment," *Wireless Personal Communications*, vol. 122, no. 3, pp. 2639–2656, Feb. 2022, doi: 10.1007/s11277-021-09022-w.

[10] T. P. Latchoumi, R. Swathi, P. Vidyasri, and K. Balamurugan, "Develop new algorithm to improve safety On WMSN in health disease monitoring," in *2022 International Mobile and Embedded Technology Conference (MECON)*, Mar. 2022, pp. 357–362, doi: 10.1109/MECON53876.2022.9752178.

[11] Y. Sahu, A. Tripathi, R. K. Gupta, P. Gautam, R. K. Pateriya, and A. Gupta, "A CNN-SVM based computer aided diagnosis of breast cancer using histogram K-means segmentation technique," *Multimedia Tools and Applications*, vol. 82, no. 9, pp. 14055–14075, Apr. 2023, doi: 10.1007/s11042-022-13807-x.

[12] Y. Qi, X. Liu, H. Li, M. Liu, and W. Li, "Discrete wavelet transform-based CNN for breast cancer classification from histopathology images," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13655 LNCS, 2023, pp. 326–340.

[13] H. U. Khan, B. Raza, M. H. Shah, S. M. Usama, P. Tiwari, and S. S. Band, "SMDetector: small mitotic detector in histopathology images using faster R-CNN with dilated convolutions in backbone model," *Biomedical Signal Processing and Control*, vol. 81, p. 104414, Mar. 2023, doi: 10.1016/j.bspc.2022.104414.

[14] Y. Alqahtani, U. Mandawkar, A. Sharma, M. N. S. Hasan, M. H. Kulkarni, and R. Sugumar, "Breast cancer pathological image classification based on the multiscale CNN squeeze model," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–11, Aug. 2022, doi: 10.1155/2022/7075408.

[15] E. M. Roopa Devi, R. Shanthakumari, R. Rajadevi, K. I. S. Shree, R. Shakkirun, and K. Poorvaja, "Imbalanced breast cancer prediction using CNN through multi-net framework," in *2023 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2023, pp. 1–7, doi: 10.1109/ICCCI56745.2023.10128543.

[16] S. Anjum *et al.*, "Lung cancer classification in histopathology images using multiresolution efficient nets," *Computational Intelligence and Neuroscience*, vol. 2023, no. 1, pp. 1–12, Jan. 2023, doi: 10.1155/2023/7282944.

[17] P. Pujar, A. Kumar, and V. Kumar, "Efficient plant leaf detection through machine learning approach based on corn leaf image classification," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, pp. 1139–1148, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp1139-1148.

[18] S. H. Sreedhara, V. Kumar, and S. Salma, "Efficient big data clustering using adhoc fuzzy c means and auto-encoder CNN," in *Lecture Notes in Networks and Systems*, vol. 563, 2023, pp. 353–368.

[19] T. P. Latchoumi, T. P. Ezhilarasi, and K. Balamurugan, "Bio-inspired weighed quantum particle swarm optimization and smooth support vector machine ensembles for identification of abnormalities in medical data," *SN Applied Sciences*, vol. 1, no. 10, p. 1137, Oct. 2019, doi: 10.1007/s42452-019-1179-8.

[20] B. Abunasser, M. R. AL-Hiealy, I. Zaqout, and S. Abu-Naser, "Convolution neural network for breast cancer detection and classification using deep learning," *Asian Pacific Journal of Cancer Prevention*, vol. 24, no. 2, pp. 531–544, Feb. 2023, doi: 10.31557/APJCP.2023.24.2.531.

[21] M. B. Tayel, M. A. A. Mokhtar, and A. F. Kishk, "Breast cancer diagnosis using histopathology and convolution neural network CNN method," in *Lecture Notes in Networks and Systems*, vol. 473, 2023, pp. 585–600.

[22] C. R. Prasad, B. Arun, S. Amulya, P. Abboju, S. Kollem, and S. Yalabaka, "Breast cancer classification using CNN with transfer learning models," in *2023 International Conference for Advancement in Technology (ICONAT)*, Jan. 2023, pp. 1–5, doi: 10.1109/ICONAT57137.2023.10080148.

[23] M. M. Zafar *et al.*, "Detection of tumour infiltrating lymphocytes in CD3 and CD8 stained histopathological images using a two-phase deep CNN," *Photodiagnosis and Photodynamic Therapy*, vol. 37, p. 102676, Mar. 2022, doi: 10.1016/j.pdpdt.2021.102676.

[24] N. Joseph and R. Gupta, "Dual multi-scale CNN for multi-layer breast cancer classification at multi-resolution," in *2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Dec. 2022, pp. 613–618, doi: 10.1109/ICAC3N56670.2022.10073985.

[25] J. G. Melekoodappattu, A. S. Dhas, B. K. Kandathil, and K. S. Adarsh, "Breast cancer detection in mammogram: combining modified CNN and texture feature based approach," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 9, pp. 11397–11406, Sep. 2023, doi: 10.1007/s12652-022-03713-3.

BIOGRAPHIES OF AUTHORS

Pattan M. D. Ali Khan earned his Bachelors of Technology B. Tech degree in CSE from JNTUA in 2005. He has obtained his Master's degree in M.E (CSE) from St. Peter's University in 2012. And currently he is a Research Scholar at B.S. Abdur Rahman Crescent Institute of Science and Technology (deemed to be university) doing his Ph.D. in Computer Science and Engineering. He has attended many workshops and induction programs conducted by various universities. His areas of interest are machine learning and image processing. He can be contacted at email: alik_12@rediffmail.com.

Dr. Xavier Arputha Rathina is an Associate Professor in Computer Science and Engineering Department at School of SCIMS of B.S. Abdur Rahman Crescent Institute of Science and Technology (deemed to be university) Chennai with an experience of 27 years in Teaching. She did her B.E in Electronics Communication Engineering and Master's degree in Computer Science and Engineering, and Ph.D. in Computer Science and Engineering. Her areas of interest are image processing, emotional intelligence, and speech processing. She can be contacted at email: xarathina@crescent.education.