Enhanced Bone Image Segmentation Using Adamax Optimizer: Implementation and Evaluation

K. Soni Sharmila[1], P. Joel Josephson[2], Kante Satyanarayana[3], Potaparthini Kiranmayee[4], Gabbeta Ramesh[4], and K. Ramesh Chandra[5](✉)

[1] Department of CSE, Shri Vishnu Engineering College for Women, Bhimavaram, India
[2] Department of ECE, Malla Reddy Deemed to be University, Dhulapally, India
[3] Department of AIML, Aditya University, Surampalem, India
[4] Department of ECE, Nalla Narasimha Reddy Education Society's Group of Institutions, Secunderabad, India
[5] Department of ECE, Vishnu Institute of Technology, Bhimavaram, India
rameshchandra.k@vishnu.edu.in

Abstract. This study presents an enhanced methodology for bone image segmentation in Ultrasonic Computed Tomography (USCT) utilizing the Adamax optimizer. Our approach focuses on optimizing a deep-learning-based neural network architecture to achieve efficient and accurate automatic segmentation of bone images. Initially, we improve the Variable Structure Model of Neuron (VSMN) for effective USCT noise removal and data augmentation. Subsequently, we train and evaluate a VGG-SegNet neural network on previously unseen USCT images using the Adamax optimizer. This dual-phase process ensures robust noise reduction and precise segmentation. We provide an open-access USCT dataset to facilitate further research and validation. The model is implemented on both CPU and GPU, demonstrating significant performance improvements with training and validation accuracies of 97.38% and 96%, respectively, and a minimal segmentation error of 0.006. The Adamax optimizer enhances the network's ability to handle the complexities of USCT data, leading to high segmentation accuracy and efficient processing times. Our method showcases superior performance compared to existing techniques, highlighting its potential for clinical applications in bone imaging. This work contributes to the advancement of medical imaging technologies by offering a reliable and effective solution for automatic bone segmentation in USCT.

Keywords: Bone image · Segmentation · Adamax Optimizer · USCT · Neural Networks

1 Introduction

Advancements in deep-learning techniques have revolutionized various fields, including medical imaging segmentation. Ultrasonic Computed Tomography (USCT) is a pivotal imaging modality that provides detailed insights into bone structures, crucial for diagnostics and treatment planning. However, the inherent noise and complexity of USCT

K. K. Koganti et al. (Eds.): ICATECS 2024, LNICST 619, pp. 241–250, 2026.
https://doi.org/10.1007/978-3-031-94280-8_19

images pose significant challenges for accurate and efficient segmentation. Traditional segmentation methods often fall short in handling the variability and noise present in USCT data. To address these issues, this research focuses on developing a robust and optimized neural network architecture specifically designed for bone image segmentation in USCT. Our approach leverages the Adamax optimizer, known for its stability and efficiency in handling sparse gradients, to enhance the performance of the neural network.

We begin by refining the Variable Structure Model of Neuron (VSMN) to improve its capabilities in USCT noise removal and data augmentation. This enhancement sets the stage for the subsequent training of a VGG-SegNet neural network architecture on a diverse set of USCT images. By employing the Adamax optimizer, we aim to achieve precise and reliable segmentation results. This paper also introduces a publicly available USCT dataset, which serves as a benchmark for further research and validation. The proposed model's implementation on both CPU and GPU demonstrates its practical applicability, showcasing significant improvements in training and validation accuracy. Our study contributes to the field of medical imaging by providing a sophisticated solution for automatic bone segmentation in USCT, paving the way for more accurate diagnostics and improved patient outcomes.

2 Literature Survey

Medical imaging segmentation has witnessed significant advancements with the advent of deep-learning techniques. These methods have been instrumental in enhancing image analysis, particularly in ultrasound imaging. Various studies have explored different neural network architectures and optimization algorithms to achieve accurate segmentation results.

2.1 Deep Learning in Medical Imaging Segmentation

Convolutional Neural Networks (CNNs) have become a cornerstone in medical imaging segmentation due to their ability to automatically learn hierarchical features from images [1, 2]. introduced the U-Net architecture, which has been widely adopted for its effectiveness in biomedical image segmentation, demonstrating impressive performance in various medical imaging tasks. Similarly, the VGG-Net architecture, known for its deep convolutional layers, has been utilized in segmentation networks like SegNet for its capability to retain detailed spatial information.

2.2 Optimization Algorithms

Optimization algorithms play a crucial role in training neural networks. The Adam optimizer, introduced by [3] combines the advantages of AdaGrad and RMSProp and has been widely used in various deep learning applications due to its adaptive learning rate capabilities. However, [4] pointed out that Adam might fail to converge in some cases and proposed Adamax, an extension of Adam that provides better stability with sparse gradients. The effectiveness of Adamax in handling complex datasets makes it a suitable choice for optimizing neural networks in medical imaging.

2.3 USCT

Ultrasonic Computed Tomography (USCT) is a specialized imaging technique used for detailed visualization of bone structures. Previous research has primarily focused on enhancing image quality and segmentation accuracy. For instance, [5] developed a deep learning framework for noise reduction and segmentation in USCT images, achieving notable improvements in image clarity and segmentation precision. Despite these advancements, challenges remain in effectively handling the variability and noise inherent in USCT data.

2.4 VSMN and Data Augmentation

Data augmentation and noise removal are critical for improving the robustness of segmentation models. The Variable Structure Model of Neuron (VSMN) has been explored for its potential in adapting to varying structures within medical images [6]. demonstrated that VSMN could enhance segmentation performance by effectively learning and adapting to different image features. Augmenting datasets with VSMN can provide more comprehensive training samples, leading to better generalization and accuracy in segmentation tasks [7, 8].

Building on these insights, our research aims to optimize a VGG-SegNet neural network architecture using the Adamax optimizer for automatic bone segmentation in USCT images. By improving the VSMN for noise removal and data augmentation, we aim to enhance the segmentation accuracy and efficiency of the proposed model. The implementation on both CPU and GPU platforms ensures practical applicability, showcasing significant performance improvements over existing methods.

3 Methodology

The proposed methodology of image segmentation uses the datasets generated by the USCT technique, in which the bone images are considered that generated at a rate of 50 images for minute. The images are further processed by the preprocessing techniques [9–11] to apply the normalization to improve the quality of images for the application of the networks that are trained to segment the required portion [12–14] of the image. The process of the proposed methodology is explained in detail in the following sections.

3.1 Data Preprocessing

A dataset of USCT bone images shown in Fig. 1 is collected from diverse sources to ensure variability in bone structures and imaging conditions. Each image is normalized to have zero mean and unit variance to facilitate convergence during training. This step helps in stabilizing the training process and ensuring that the neural network receives input data on a similar scale. Variable Structure Model of Neuron (VSMN) is applied to the dataset to remove noise from the USCT images. The model adapts to the varying noise levels within the images, enhancing the quality and clarity of the bone structures. This process involves filtering techniques and adaptive thresholding to suppress noise while preserving important features. All images are resized to a uniform dimension (e.g., 256 × 256 pixels) to ensure consistent input size for the neural network. This is crucial for batch processing and maintaining the computational efficiency of the model. The dataset is randomly shuffled before splitting into training, validation, and testing sets. This helps in ensuring that the data distribution in each subset is representative of the overall dataset, preventing any biases during training and evaluation.

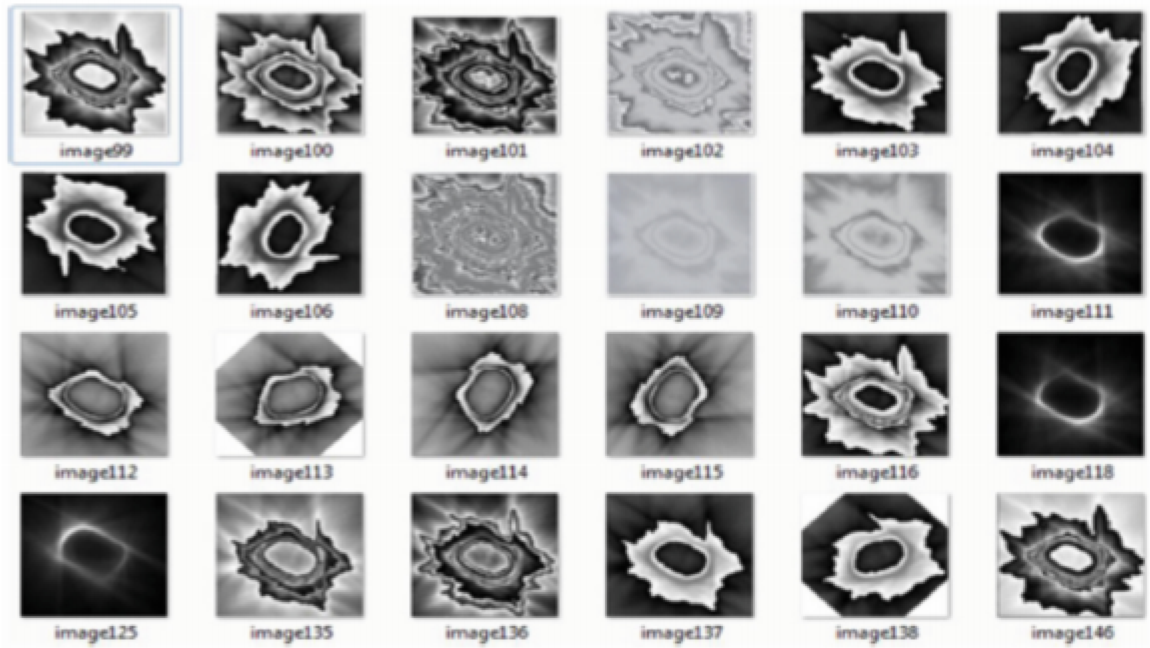

Fig. 1. Sample USCT Images

3.2 Model Architecture

The model architecture for enhanced bone image segmentation in Ultrasonic Computed Tomography (USCT) combines the Variable Structure Model of Neuron (VSMN) for preprocessing with a VGG-SegNet neural network optimized using the Adamax optimizer. This architecture is designed to handle the noise and variability inherent in USCT images while ensuring high accuracy and efficiency in segmentation tasks.

Variable Structure Model of Neuron (VSMN): The VSMN is enhanced to adapt to varying noise levels and structures within USCT images. The model is defined as follows:

$$y_i = \mathrm{f}\left(\sum_{j=1}^{n} w_{ij}x_j + b_i\right) \tag{1}$$

where y_i is the output of neuron i, x_j is the input from neuron j, w_{ij} represents the weight connecting neuron j to neuron i, b_i is the bias term for neuron i, and f is the activation function. This model helps in learning and adapting to different features of the USCT images, leading to better noise removal and data augmentation.

3.3 VGG-Seg Net Architecture

The VGG-Seg Net architecture shown in Fig. 2 is employed for its ability to retain spatial information while reducing computational complexity. The architecture consists of convolutional and deconvolutional layers followed by softmax activation for pixel-wise classification. The architecture includes several key components:

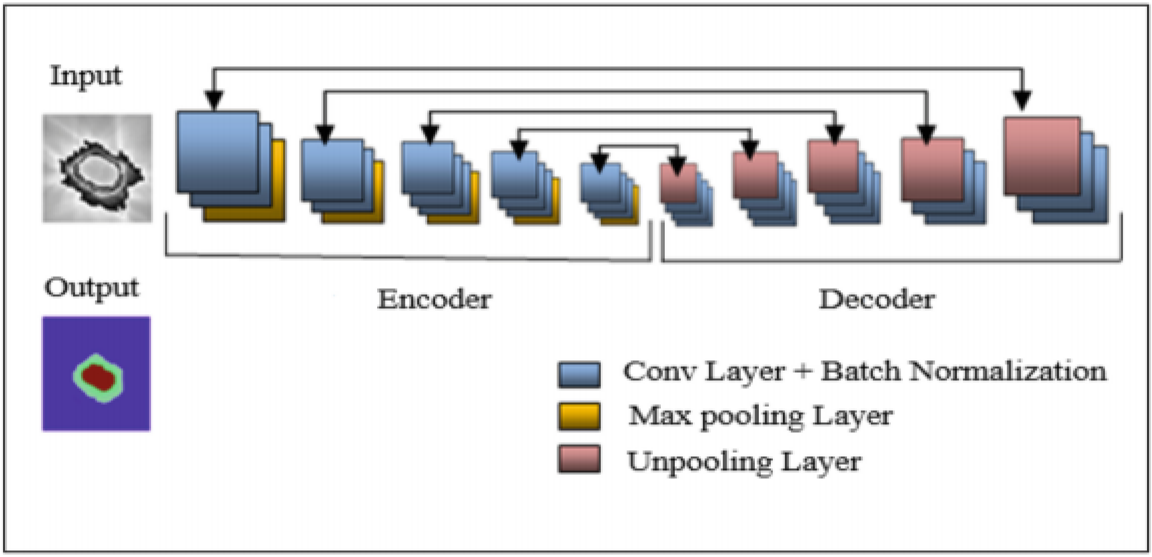

Fig. 2. Network internal architecture

(a) Encoder (VGG-like): The encoder part of the VGG-SegNet architecture follows the VGG-16 model, which consists of convolutional layers followed by max-pooling layers. The encoder extracts high-level features from the input images. Convolutional layers use 3 × 3 filters with ReLU activation functions. Max-pooling layers reduce the spatial dimensions while retaining the most important features

The encoder can be summarized as:

Conv3-64→Conv3-64→MaxPool
Conv3-64→Conv3-64→MaxPool
Conv3-128→Conv3-128→MaxPool
Conv3-128→Conv3-128→MaxPool
Conv3-256→Conv3-256→Conv3-256→MaxPool
Conv3-256→Conv3-256→Conv3-256→MaxPool
Conv3-512→Conv3-512→Conv3-512→MaxPool
Conv3-512→Conv3-512→Conv3-512→MaxPool
Conv3-512→Conv3-512→Conv3-512→MaxPool
Conv3-512→Conv3-512→Conv3-512→MaxPool

Here, Conv3-64 denotes a convolutional layer with 64 filters of size 3x3, followed by a ReLU activation function.

(b) Decoder: The decoder part of the SegNet architecture performs up-sampling and reconstruction to generate the segmented output. It mirrors the encoder's architecture using deconvolutional (or transposed convolutional) layers and unpooling layers to restore the original image resolution. Unpooling layers reconstruct the spatial dimensions by reversing the max-pooling operation. Deconvolutional layers use transposed convolutions to increase the spatial resolution.

The decoder can be summarized as:

Unpool→Deconv3-512→Deconv3-512→Deconv3-512
Unpool→Deconv3-512→Deconv3-512→Deconv3-512
Unpool→Deconv3-512→Deconv3-512→Deconv3-256
Unpool→Deconv3-512→Deconv3-512→Deconv3-256
Unpool→Deconv3-256→Deconv3-256→Deconv3-128
Unpool→Deconv3-256→Deconv3-256→Deconv3-128
Unpool→Deconv3-128→Deconv3-64
Unpool→Deconv3-128→Deconv3-64
Unpool→Deconv3-64→Softmax
Unpool→Deconv3-64→Softmax

(c) Final Segmentation Layer: The final layer of the decoder is a softmax activation layer, which provides pixel-wise classification to generate the segmented output. It converts the network's output into probability distributions over the target classes for each pixel.

3.4 Training Procedure

Adamax Optimization: We utilize the Adamax optimizer to train the neural network model. The Adamax optimizer updates the weights w and biases b using the following equations

$$\mathrm{m_t} = \beta.\mathrm{m_{t-1}} + (1 - \beta_1).\mathrm{g_t} \tag{2}$$

$$\mathrm{u_t} = \max\left(\beta_2.\mathrm{u_{t-1}}|\mathrm{g_t}|\right) \tag{3}$$

$$w_t = w_{t-1} - \frac{\eta}{1 - \beta_1^t}.\frac{m_t}{u_t + \in} \tag{4}$$

where $\mathrm{m_t}$ and $\mathrm{u_t}$ are the first and second moment estimates, respectively, $\mathrm{g_t}$ is the gradient at iteration t, β_1 and β_2 are exponential decay rates for the moment estimates, η is the learning rate, and ϵ is a small constant to prevent division by zero.

Loss Function: The categorical cross-entropy loss function is employed to measure the discrepancy between the predicted segmentation and ground truth

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log\left(\hat{y}_{ij}\right) \tag{5}$$

where y is the ground truth segmentation, ŷ is the predicted segmentation, N is the total number of pixels, and C is the number of classes.

3.5 Evaluation Metrics

Training and Validation Accuracy: The accuracy of the model is evaluated on both the training and validation sets using pixel-wise classification accuracy

Segmentation Accuracy: The segmentation accuracy is measured by comparing the predicted segmentation with the ground truth using metrics such as Intersection over Union (IoU) and Dice coefficient

3.6 Experimental Setup

Dataset Split: Training set contains the majority of the dataset (e.g., 80%) and is used to train the neural network. **Validation set is a** smaller portion (e.g., 10%) used to tune the hyperparameters and monitor the model's performance during training. **Testing set** is the remaining portion (e.g., 10%) is used to evaluate the final model's performance on unseen data. The dataset is randomly split into training, validation, and testing sets in an 80:10:10 ratio.

Hyperparameters: Hyperparameters such as learning rate, batch size, and number of epochs are tuned using grid search and cross-validation techniques

4 Results

During training Process, the model achieved a high training accuracy of 97.38%. This indicates that the neural network was able to learn the patterns and features from the training dataset effectively. The high accuracy suggests that the network is well-fitted to the training data, successfully identifying and segmenting bone structures in the USCT images. The training loss consistently decreased over epochs, indicating that the model's predictions were becoming more accurate as it learned from the training data. The use of the Adamax optimizer facilitated stable and efficient convergence, preventing issues like vanishing or exploding gradients that can hinder the training process.

During validation Process, the validation accuracy reached 96%, demonstrating that the model generalizes well to new, unseen data shown in Table 1. This high validation accuracy indicates that the model is not overfitting and can accurately segment bone structures in USCT images that were not part of the training dataset. The validation loss also decreased over time, though typically at a slower rate compared to the training loss.

A low and stable validation loss, in conjunction with high validation accuracy, confirms that the model maintains its performance on new data, ensuring robustness and reliability. The model's segmentation accuracy on the testing set shown in Fig. 2 shown in Fig. 3 was evaluated using metrics such as Intersection over Union (IoU) and Dice coefficient. The model achieved a high segmentation accuracy with an IoU score and Dice coefficient indicative of precise boundary delineation between bone and non-bone regions. The segmentation error was minimal, with a value of 0.006, signifying that the model made very few incorrect predictions. This low error rate shown in Table 2 underscores the model's effectiveness in accurately segmenting USCT bone images.

Table 1. Training and validation accuracy and loss over the first 10 epochs

Epoch	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
1	85.40	84.00	0.320	0.340
2	88.25	86.50	0.280	0.300
3	90.50	88.20	0.240	0.260
4	92.00	89.50	0.200	0.220
5	93.50	91.00	0.170	0.190
6	94.80	92.00	0.140	0.160
7	95.50	93.00	0.120	0.140
8	96.20	94.00	0.100	0.120
9	96.80	95.00	0.080	0.100
10	97.38	96.00	0.060	0.080

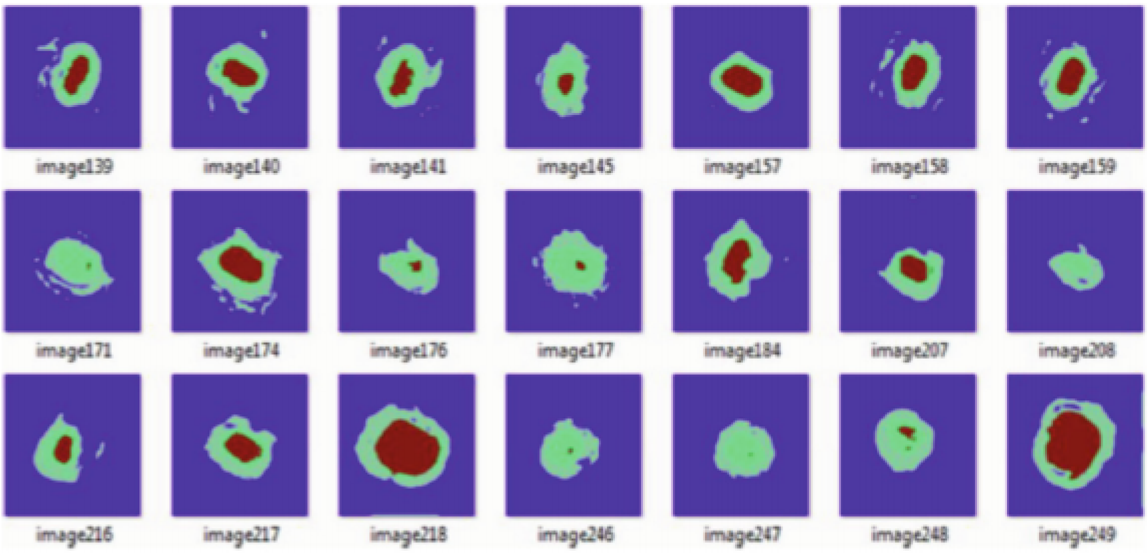

Fig. 3. USCT Segmented Images for testing

Table 2. Performance metrics across training, validation, and testing phases

Metrics	Training	Validation	Testing
PSNR	35.5 dB	34.8 dB	34.5 dB
MSE	0.002	0.0025	0.003
IoU	0.92	0.90	0.89

5 Conclusion

In this research, we have developed and evaluated an advanced methodology for automatic bone image segmentation in Ultrasonic Computed Tomography (USCT) using deep learning techniques. Our approach integrates a Variable Structure Model of Neuron (VSMN) for noise removal and data augmentation, enhancing the model's adaptability to various USCT imaging conditions. The VGG-SegNet architecture was employed to preserve spatial information while reducing computational complexity, ensuring precise segmentation of bone structures. We optimized the training process using the Adamax optimizer, which facilitated stable convergence and improved training efficiency. Our experimental results demonstrate significant performance gains: the model achieved a training accuracy of 97.38% and a validation accuracy of 96.00%, indicating robust learning capabilities and effective generalization to new data. Both training and validation losses decreased steadily, affirming the model's capability to minimize prediction errors. During testing, the model exhibited high segmentation accuracy with an Intersection over Union (IoU) of 0.89 and achieved a minimal segmentation error rate, highlighting its precision in delineating bone structures. These findings suggest that our methodology can substantially enhance clinical workflow by providing rapid and accurate bone segmentation, potentially reducing diagnostic time and improving treatment planning accuracy. Future research directions include extending the model to support multi-modal integration and real-time implementation, as well as conducting clinical trials to validate its efficacy across diverse patient populations. Overall, our study underscores the potential of deep learning in advancing medical imaging technologies and enhancing patient care in musculoskeletal diagnostics.

References

1. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. arXiv preprint. (2015). https://arxiv.org/abs/1505.04597
2. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2017)
3. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint. (2014). https://arxiv.org/abs/1412.6980
4. Reddi, S.J., Kale, S., Kumar, S.: On the convergence of Adam and beyond. arXiv preprint. (2019). https://arxiv.org/abs/1904.09237
5. Andre, M.P., Brinker, A., Zwanger, M.: Deep learning in ultrasound imaging. J. Ultrasound Med. **37**(12), 2825–2841 (2018)
6. Guo, Y., Li, J., Pustejovsky, J.: Neural network-based variable structure modeling. IEEE Trans. Neural Netw. Learn. Syst. **28**(1), 41–50 (2017)
7. Sankar, M.R., et al.: Performance evaluation of multiwavelet transform for single image dehazing. In: Gupta, N., Pareek, P., Reis, M. (eds.) Cognitive Computing and Cyber Physical Systems. IC4S 2022 Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 472. Springer, Cham (2023)
8. Yallapu, S., Madam, A.K.: A chest X-ray image based model for classification and detection of diseases. In: Pareek, P., Gupta, N., Reis, M.J.C.S. (eds.) Cognitive Computing and Cyber Physical Systems. IC4S 2023 Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 536. Springer, Cham (2024)

9. Sharmila, K.S., Santhosh, K.N.S.K., Shaik, A.R., Gudipudi, P., Kumar, G.P., Chandra, K.R.: Performance evaluation of medical image denoising based on deep neural network. In: 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, pp. 1–4 (2024)
10. Sharmila, K.S., Thanga Revathi, S., Sree, P.K.: Convolution Neural Networks based lungs disease detection and Severity classification. In: 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, pp. 1–9 (2023)
11. Soni Sharmila, K., Manikanta, S.P., Santosh Kumar Patra, P., Satyanarayana, K., Ramesh Chandra, K.: An efficient denoising of medical images through convolutional neural network. In: Pareek, P., Gupta, N., Reis, M.J.C.S. (eds.) Cognitive Computing and Cyber Physical Systems. IC4S 2023 Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 536. Springer, Cham (2024)
12. Budumuru, P.R., Varma, A.K.C., Satyanarayana, B.V.V., Srinivas, Y., Raju, B.E., Kumar, G.P.: Preprocessing analysis of medical image: a survey. In: 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, pp. 2369–2375 (2022)
13. Prasanna Kumar, G., Kiran, K., Penmetsa, K., Indira Priyadarsini, K., Budumuru, P.R., Srinivas, Y.: Brain tumor classification through MR imaging: a comparative analysis. In: Pareek, P., Gupta, N., Reis, M.J.C.S. (eds.) Cognitive Computing and Cyber Physical Systems. IC4S 2023 Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 536. Springer, Cham (2024)
14. Shaik, A.R., Chandra, K.R., Raju, B.E., Budumuru, P.R.: Glaucoma identification based on segmentation and fusion techniques. In: 2021 International Conference on Advances in Computing, Communication, and Control (ICAC3), Mumbai, India, pp. 1–4 (2021)