

DUALATT-VITRESNET WITH FENNEC FOX OPTIMIZED HYPER PARAMETERS FOR AUTOMATED OSTEOPOROSIS CLASSIFICATION USING MEDICAL IMAGING

V. Venkatachalam^{*1}, N. Subash², P. Archana Devi³, Saranya S⁴, R. Tharwin Kumar⁵,
Chokkan Priyadharshini⁶

¹Department of Computer Science and Engineering, Erode Sengunthar Engineering College, Thudupathi, Perundurai, Erode, Tamilnadu, India.

²Department of Electronics and Communication Engineering, Roever Engineering College, Perambalur, Tamilnadu, India.

³Department of Information Technology, Mailam Engineering College, Tamilnadu, India.

⁴Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India.

⁵Department of Electrical and Electronics Engineering, Mohan Babu University, Tirupati, Andhra Pradesh, India.

⁶Department of Electronics and Communication Engineering, Bharathiyar Institute of Engineering for Women, Salem, Tamilnadu, India.

¹<https://orcid.org/0009-0003-7191-8337>, ²<https://orcid.org/0009-0006-8127-4282>, ³<https://orcid.org/0009-0004-1436-2480>,

⁴<https://orcid.org/0009-0004-9300-9078>, ⁵<https://orcid.org/0000-0002-3438-4562>, ⁶<https://orcid.org/0009-0006-4160-4742>,

Email: profdrv@gmail.com, drsubash777@gmail.com, archana.devi1995@gmail.com, saranyas.work@gmail.com, tharwin.eee@gmail.com, priyadharshinii1988@gmail.com

ARTICLE INFO

Article History

Received: January 30, 2026

Reviewed: April 1, 2026

Accepted: May 1, 2026

Published: May 29, 2026

Keywords:

Gaussian filter,
Style GAN,
DenseNet121,
DualAtt-ViTResNet,
FFOA.

ABSTRACT

Osteoporosis is a progressive bone disease which is necessary to reduce fracture risk and complications. This paper aims to establish an automated classification system for osteoporosis based on automated image processing of medical images, and to generate a sophisticated method of combining different processing techniques, such as pre-processing, data augmentation, feature extraction, and classification modelling using advanced Deep Learning (DL) methods. Preprocessing of images involves resizing and the use of a Gaussian filter to eliminate noise from the image and thus improve image quality. Augmentation of images occurs through the use of Style GAN to create synthetic images, reduce the effects of data imbalance and improve the generalization of a model trained on limited sample sizes. Feature extraction occurs through the use of DenseNet121 which is demonstrated to be very efficient in producing a large number of high-quality hierarchical features. The techniques for classification of the extracted features are based on a novel Dual-Attention Vision Transformer and ResNet architecture (DualAtt-ViTResNet). It provides strong local feature representation of ResNet and the ability to acquire contextual information at the global level through vision transformers and their dual attention mechanism. Finally, hyper-parameter optimization of the DualAtt-ViTResNet architecture is accomplished using the Fennec Fox Optimization Algorithm (FFOA) which is used to improve accuracy, speed, and robustness of the classification task. A detailed experimental evaluation of the proposed framework indicates that the automated osteoporosis classification framework outperformed the existing frameworks with respect to higher accuracy of 0.9744, precision of 0.9767, Recall of 0.9789 and F1-Score of 0.9777. The proposed system offers an efficient solution for automated osteoporosis detection, supporting clinical decision-making and early diagnosis in medical imaging applications.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

A reduction in Bone Mineral Density (BMD) and the breakdown of the microstructure of bone tissue are two characteristics that define Osteoporosis. This causes the bone to become weak and more likely to fracture as a result of the increased amount of energy placed on the bones [1]. Currently estimated to be approximately 23.1% of women and 11.7% of men worldwide are affected by osteoporosis according to a meta-analysis of studies conducted between 2020 and 2021.

There is also a considerable variation in the rates of osteoporosis between individual countries around the world [2], [3]. Fragility fractures occur as a result of the normal biomechanics of our bodies being altered. These fractures frequently lead to higher rates of mortality, as well as the need for additional medical treatment, rehabilitation, and prolonged hospital stays and have been well documented in previous research to have major physical, psychological, social, and economic consequences [4]. These consequences include reduced quality of life, increased costs associated with treating osteoporosis, increased risks of mortality, decreased levels of physical activity, and decreased independence among elderly individuals [5]. As a result of the potentially life-threatening complications that occur from these osteoporotic fractures, it is vital to detect osteoporosis as early as possible so that appropriate measures are taken to prevent the occurrence of osteoporosis and the associated complications arising from an osteoporotic fracture including pain and suffering. For this reason, the World Health Organization (WHO) has deemed Dual Energy X-ray Absorptiometry (DEXA) to be the standard for assessing BMD and diagnosing osteoporosis [6], [7]. The BMD values that DEXA provides are expressed in terms of T-scores, which are calculated by taking the difference between an individual's measured BMD and the mean BMD for a reference population.

Osteoporosis is defined as having a T-score less than -2.5 and osteopenia has a T-score between -1.0 and -2.5. DEXA is still the most prevalent quantitative method of evaluating bone density due to its ease of access through the skin, affordability and limited health risks. However, to be truly effective and reliable, it must be viewed in conjunction with other techniques to produce meaningful outcomes. DEXA is limited to BMD, with low screening rates reported [8-10]. Additionally, the BMD results from DEXA are based on a 2-dimensional plane, which limits the ability to accurately calculate BMD and predict fracture risk. Additionally, DEXA results are skewed by natural age-related changes to bone strength, which could lead to a miscalculation of BMD and also caused by outside forces affecting the bone [11]. To mitigate these limitations, the proposed system uses a hybrid model to classify osteoporosis. Using medical imaging, this research establishes an automated framework for classifying osteoporosis. The framework is based on the DualAtt-ViTResNet architecture and employs FFO for hyperparameter tuning. This research includes numerous process for the identification of osteoporosis classification [12]. The process initiates with input medical images, which goes through preprocessing steps to remove noise and to standardize image quality. Thereby to accomplish this, a Wiener filter is used to reduce the mean square error of the original image compared to the restored image based on both local image variance and the statistical characteristics of the noise, thus providing a means for preserving the most important bone structure and intensity variations from image to image.

However, because of reliance on accurate estimates of the signal and noise statistics, a Wiener filter's effectiveness are limited by incorrect assumptions, causing residual noise or an overly smooth appearance of the restored image [13-15]. Hence, the proposed system uses Gaussian filter to resolve this shortcomings. In order to improve the resilience of the model, data augmentation is also carried out, which helps produce a variety of permutations of the input image. The preprocessed image is then fed to feature extraction phase which provides dense representations of spatial and temporal endurance in the input images. The multi-scale feature extraction capabilities of Inception-V3 is used to detect small differences in texture and density of bones in medical images [16]. The depth and complexity of the model also create challenges in terms of the computational power and memory needed to use the model, therefore it is difficult to use the model in resource-limited medical environments. ResNet-50 architecture enables the extraction of robust and representative features of bone density and texture combined [17-18]. As a result, this form of analysis provides superior prediction accuracy for classifying bone tissue types. This high-level feature extraction method therefore requires a large number of parameters and high calculations and long processing times due to the associated resource demand of the model. To address this issue, DenseNet-121 model is used in the proposed system [19-20]. The extracted image is passed to the classification stage to enhance the accuracy of the class predictions by concentrating on critical portions of the image. Comparison of various existing classifications are depicted in Table 1.

Table 1: Related works for classification models to detect Osteoporosis.

Sl.N O	Author/Year	Classification	Merits	Demerits	Performance metrics
1.	Mukhtar Ul Islam <i>et al</i> [2025] [21]	ViT-16	The Attention Mechanism in the ViT-16 enables it to sift through scans and apply weight to different areas of the scan.	It requires careful considerations of the associated computational costs and available data.	It attains an accuracy of 88%, precision of 85%, Recall of 88% and FI-score of 86%.
2.	M. J. Aashik Rasool <i>et al</i> [2024] [22]	KONet	KONet uses an ensemble approach being more stable and decreasing the chances of being biased.	The architecture is more complicated and it is more challenging to deploy & maintain in a clinical setting.	The performance metrics of KONet is 95% of accuracy, precision, recall and FI-Score.
3.	Abdul Wahab Muzaffar <i>et al</i> [2025] [23]	OsteoNet	OsteoNet performs well with moderate amounts of data, particularly when utilizing data augmentation and transfer learning techniques.	It struggle to perform well with respect to long-range dependencies and global contextual information in an image.	It achieves an accuracy of 74.1% and FI-Score of 73.9%
4.	Sailaja Thota <i>et al</i> [2024] [24]	MobileNet	It's computational efficient model utilises a depth wise separable convolutional network that significantly reduces model size and provides a lower computational requirement.	Due to its compact nature, it is not able to extract the detailed features that are required to accurately diagnose osteoporosis.	MobileNet attains an accuracy of 94%, precision of 91% and recall of 92%.
5.	Athira O M <i>et al</i> [2025] [25]	ResNet-50+GRU	It reduces the need to collect large amounts of data specifically for osteoporosis	It is computationally expensive and requires additional memory	It establishes an accuracy of 95.65%, precision of 95.32%, recall of 95.98% and FI-Score of 95.49%.

Source: Authors, (2026).

Existing osteoporosis classification methods have difficulties in extracting features effectively and have very low classification accuracies for bone imaging features analyzed with medical images. Therefore to overcome these issues, DualAtt-ViTResNet model is used in the proposed system to analyze the developed features of bones. Furthermore, to enhance the classification stage, FFOA is implemented to provide significant improvement in robustness, accuracy, and computational efficiency to achieve success at automating osteoporosis classification in the clinical imaging environment.

I.1 OBJECTIVES

The main objectives of this research includes,

- To reduce noise and to enhance medical image quality, preprocessing techniques such as resizing and Gaussian filter is implemented.
- To address imbalanced datasets and to increase data diversity and robustness, data augmentation method, StyleGAN is used.
- To extract discriminative and significant features from medical images and to effectively represent bone functions, DenseNet-121 is executed.
- To incorporate both channel-wise and spatially significant features, DualAtt-ViTResNet is employed.
- To increase model efficiency and classification accuracy, FFOA is introduced.

I.2 PAPER ORGANIZATION

This paper is organized as follows: Section II provides an explanation of the suggested system description. The study methodology is thoroughly described in Section III, which also includes examples of the process, methods, and mathematical representations. The experimental results of this work are described in Section IV, which also include the simulation setup, comparison analysis, and research summary. The conclusion of the proposed investigation is explained in Section V.

II. PROPOSED SYSTEM DESCRIPTION

The proposed framework for image classification includes preprocessing, augmentation, feature extraction, classification and hyperparameter tuning. The initial step in this process involves resizing and denoising with a Gaussian filter, ensuring it maintain consistent quality throughout the process. After denoising, StyleGAN is used for augmentation technique to create multiple realistic synthetic images that help to improve generalization. Once all synthetically generated images are created, the original images and all the synthetically generated images are run through DenseNet-121 to extract deep features from all of the images. This step enables to acquire effective spatial and contextual representations from the images efficiently. The input from the deep features generated from the images are passed into a DualAtt-ViTResNet classifier. The DualAtt-ViTResNet classifier uses convolutional and transformer types of learning, along with dual attention mechanisms, to capture local details and global dependencies simultaneously. Figure 1 illustrates the block diagram of the proposed Fennec Fox Optimized DualAtt-ViTResNet for automated osteoporosis classification.

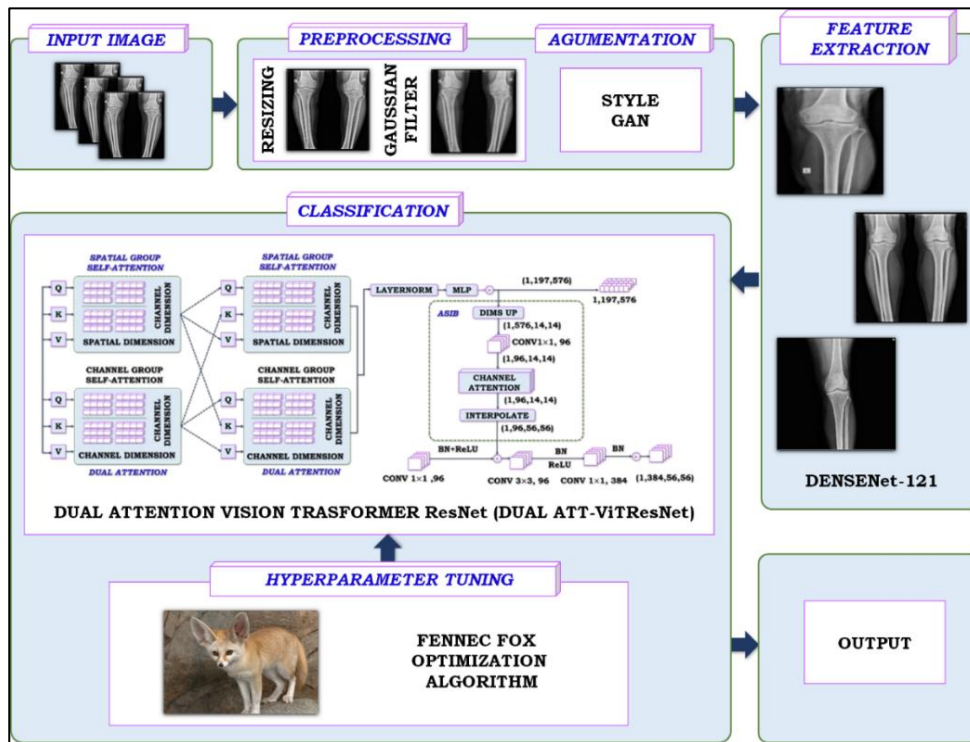


Figure 1: Block diagram of the proposed model.

Source: Authors, (2026).

Finally, in order to further enhance the model's performance, the FFOA is utilized for hyperparameter tuning. This algorithm ensures that the best possible model configuration are produced for classifying images accurately and consistently.

III. PROPOSED SYSTEM MODELLING

III.1 PREPROCESSING

III.1.1 Resizing

By using standardised medical images at a defined resolution prior to analysis through resizing, DL models achieve a uniform input. Resized medical images are less computationally expensive compared to other forms of medical imaging, and also maintain all important bone structural characteristics, thereby making it easier for neural networks to extract features related to the classification of osteoporosis and perform the correct classification of osteoporosis.

III.1.2 Gaussian Filter

The process of classifying osteoporosis with an automated method requires the preprocessing of digital images using computer vision such as X-ray and Dual Energy Absorptiometry [DXA]. Medical images typically have noise that obscure the natural texture of bone and therefore negatively impact the diagnosis of osteoporosis, thus to eliminate any noise or blurring from the image prior to exposure to feature extraction and classification, a Gaussian filter is used. Gaussian smoothing is performed by performing convolution with a Gaussian operator, which is effective in reducing the presence of high frequency noise, therefore allowing for the preservation of bone structure. The one-dimensional Gaussian filter is scientifically stated as,

$$G_{1D}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (1)$$

For medical images, the 2-D symmetrical homogeneous Gaussian filter is described as,

$$G_{2D}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

The Gaussian distribution function contains a variable σ , is the Standard deviation, which directly controls the level of smoothing produced by the filter. The filtering process is accomplished by convolving the pixel value matrix of the original image with a specific normalized Gaussian Kernel. Using the following Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), the quality of the filtered bone image, is used to classify Osteoporosis with high degree of accuracy, and is measured by comparing the MSE of the original Image with the filtered image of size $(M \times N)$.

$$MSE = \frac{1}{M \times N} \sum_{i,j} (O_{image} - R_{image})^2 \quad (3)$$

PSNR quantifies the enhancement of the quality of an image that results from the filtering process.

$$PSNR = 10 \log_{10} \left(\frac{255 \times 255}{MSE} \right) \quad (4)$$

Gaussian filters increase the visibility of the structural and textural patterns associated with bone, and subsequently provide more accurate, reliable, and effective automated classification of osteoporosis by reducing noise and improving the PSNR of the image being analysed.

III.2 AUGMENTATION- STYLE GAN

StyleGAN allows to generate the synthetic images of bone using the principles of free-form image creation. Through this use of StyleGAN, many of the issues related to the small size and imbalance of the medical image databases used for automated classification of osteoporosis are addressed. In order to generate images using StyleGAN, first select random latent vectors $z \in Z$ from a two-dimensional normal distribution. These latent vectors are transformed into a three-dimensional space (W) using a learned transformation from the four-dimensional style space. The three-dimensional representation are then transformed into Style Space (S), where each of the individual style channels controls certain aspects of the generated image such as the texture of the bone, the density of the bone and the arrangement of the trabecular and cortical bone. The use of StyleGAN's Style Space allows for the generation of synthetic images that have anatomical accuracy and diversity compared to the currently available medical images. StyleGAN's Style Space is also unscrambled, allowing for the precise control of clinically relevant features, such as the arrangement of the trabecular bone and the thickness of the cortex during the generation of synthetic images. The availability of anatomically accurate and diverse synthetic images enhances the pool of images available for training DL systems. Thus, this enhanced training pool increases the feature learning of the DL systems and increases the accuracy and robustness of the automated classification of osteoporosis.

III.3 FEATURE EXTRACTION-DENSENET-121

DenseNet121 attains exceptional performance in automatically classifying osteoporosis through its dense connectivity, which allows for all layers of the model to receive the same input feature maps from every previous layer. With this architecture, DenseNet121 easily reuse many of the features used by the previous layers of the network, which in turn required for accurately detecting even slightly different bone densities or structures. The architecture of DenseNet-121 is illustrated in Figure 2.

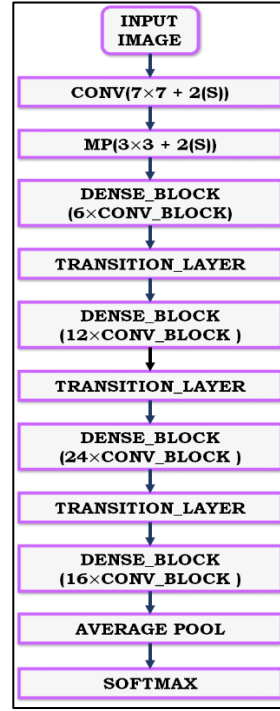


Figure 2: Structure of DenseNet-121.

Source: Authors, (2026).

Mathematically, the output from the n^{th} layer is expressed as follows,

$$Y_n = C_n([Y_1, Y_1, \dots, Y_{n-1}]) \quad (5)$$

Where, C_n is a composite function of convolution, batch normalization, and ReLU activation, and $[Y_1, Y_1, \dots, Y_{n-1}]$ are concatenated inputs from all previous feature maps in the network. Additionally, DenseNet121 has incorporated bottleneck layers to improve the computation performance when training, this layer utilizes a 1×1 convolution followed by a 3×3 convolution defined as,

$$Y_{bottleneck} = W_{3 \times 3} \left(\text{ReLU} \left(\text{BN} \left(W_{1 \times 1} (Y_n) \right) \right) \right) \quad (6)$$

Where $W_{1 \times 1}$ and $W_{3 \times 3}$ are weight matrices for the 1×1 and 3×3 convolutions, respectively, and BN is the batch normalization. Transition layers are composed of a 1×1 convolution and a 2×2 average pooling operation, which decrease the spatial dimensions of the feature map, the transition layer is defined as follows,

$$Y_{trans} = \text{Pooling} \left(\text{ReLU} \left(\text{BN} \left(W_{1 \times 1} (Y_n) \right) \right) \right) \quad (7)$$

DenseNet121 possess the added benefit of using its ability to concatenate feature maps from all prior layers, allowing it to efficiently learn both fine-grained trabecular bone structures and macro-level skeletal structures which are important in correctly identifying bones as normal, osteopenic, and osteoporotic. Additionally, due to the dense connectivity of DenseNet121, the number of weights are substantially reduced and therefore reduces overfitting on smaller osteoporosis datasets while achieving high classification accuracy.

III.4 CLASSIFICATION- DUALATT-VITRESNET

For automated classification of osteoporotic bone, fine-grain local image structures and skeleton context need to be identified at a global level. Currently, most approaches utilise the transformer model of Vision Transformers, by captivating image input, breaking it into $P \times C$ patches and reconstructing them with added positional encodings. Inappropriately, this process of patching reduce the inherent 2D layout of bone images, this reduced layout makes it difficult to detect changes due to osteoporosis in both trabecular and cortical bone. Therefore, to provide a solution, a dual attention mechanism is utilised that operates in parallel for both the spatial and channel dimensions. In terms of layout structure, the feature map was grouped by using N_g groups where each group contained P_g number of patches. Within each group self-attention was computed independently, therefore collectively creating an attention map for each group. This combined map is referred to as Grouped Spatial Attention, it is formulated as,

$$A_s(Q, K, V) = \text{Concat} \left(A_1^*(Q_1, K_1, V_1), \dots, A_{N_g}^*(Q_{N_g}, K_{N_g}, V_{N_g}) \right) \quad (8)$$

The local spatial groups for obtaining Q, K, V through a linear combination of the input features and $Q_i, K_i, V_i \in R^{P_g \times C}$, local attention operation retain local spatial information, including instances of trabecular thinning and deterioration of micro texture. At the same time, channel-wise self-attention was utilized on the channel dimension ruined up into N_w groups, each containing a total of C_w channels so that $C = C_w \times N_w$.

$$A_c(Q, K, V) = \text{Concat}(A_1^*(Q_1, K_1, V_1), \dots, A_{N_w}^*(Q_{N_w}, K_{N_w}, V_{N_w})) \quad (9)$$

This allows for the overall density of bone and its integrity to be represented in channel-level attention when all spatial locations of the image together are considered at the same time. Cross-fusion is the integrated strategy of ensuring the effective exchange of information between local and global representations during the processing of both types of attention, which allows channel attention to influence discriminative local regions which are selected for the channel branch of the neural network. The diagram of DualAtt-ViTResNet is depicted in Figure 3.

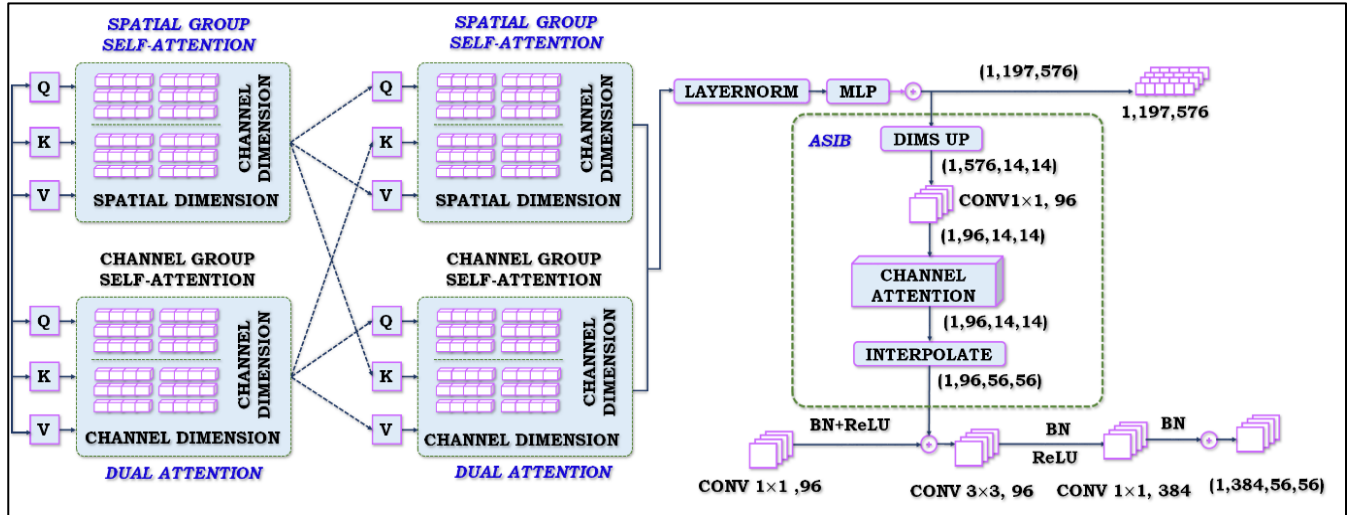


Figure 3: Structure of DualAtt-ViTResNet.

Source: Authors, (2026).

In the DualAtt-ViTResNet, the ViT part uses the initial layers of ResNet to make the feature maps into patch features that are then flattened into 196 patch features $x_p = \{x_p^1, x_p^1, \dots, x_p^{196}\}$ to build an embedded sequence including the position information. The embedded sequence is built by,

$$z = [x_p^1, x_p^2, \dots, x_p^{196}] + E_{pos} \quad (10)$$

Where, E is the patch embedding matrix and E_{pos} is the position embedding. To classify the images, the sequence is given a learnable class symbol x_{class} to create a sequence,

$$z_0 = [x_{class}, x_p^1, x_p^2, \dots, x_p^{196}] + E_{pos} \quad (11)$$

This sequence captures the full context of the entire bone image, providing a complete view of the overall structure of the skeleton but losing local spatial relationships. The multi-head self-attention is applied to the embedded sequence where each head computes and the outputs are concatenated and linearly projected to obtain,

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (12)$$

$$Multi\ Head(Q, K, V) = Concat(Head_1, \dots, Head_9)W^0 \quad (13)$$

The ViT branch of this approach allows for long-range dependency capture between different areas of the bone. The proposed technique incorporates patch flattening to create a disconnection between the area, and this disconnection is compensated through the grouped spatial attention mechanism, while the grouped channel attention allows for the maintenance of overall global semantic consistency in bone representation. By combining the ViT and ResNet architectures, the DualAtt-ViTResNet leverage patch embedding with positional encoding, grouped spatial self-attention, grouped channel self-attention, and multi-head self-attention to create a well-balanced representation of the local microstructure of a bone in conjunction with the full bone setting to significantly improve feature discriminability and robustness, thus rendering it mainly as an automated system for potentially accurate and reliable classification of osteoporosis.

III.5 FENNEC FOX OPTIMIZATION ALGORITHM

Automated classification of Osteoporosis is expressed as an optimization problem where the objective is to find the optimal parameters of the DL algorithm that yield the best classification performance for Osteoporosis.

The FFOA uses the fennec fox as a representation of a possible solution. The position of the i^{th} fennec fox is represented by a vector dimension m , where m is the number of decision variables for Osteoporosis classification, denoted as $X_i = (x_{i,1}, x_{i,2}, \dots, x_{i,m})$. Therefore, the initial population of candidate solutions is randomly populated throughout the search space to promote a diverse population and allow the algorithm to investigate multiple sets of parameter configurations.

$$x_{ij} = lb_j + r. (ub_j - lb_j) \quad (14)$$

The candidate solutions are stored in a population matrix where each row corresponds to a potential solution for Osteoporosis classification and each column corresponds to a unique decision variable.

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} x_{1,1} & \dots & x_{1,j} & \dots & x_{1,m} \\ \vdots & & \vdots & & \vdots \\ x_{i,1} & \dots & x_{i,j} & \dots & x_{i,m} \\ \vdots & & \vdots & & \vdots \\ x_{N,1} & \dots & x_{N,j} & \dots & x_{N,m} \end{bmatrix} \quad (15)$$

Once the population has been initialized, the algorithm evaluates each solution by taking the set of candidate values and using them as input to the optimization objective function to obtain the corresponding fitness values.

$$F = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix} \quad (16)$$

These fitness metrics define the classification performance in terms of sensitivity to identify the Osteoporosis classification label of the subject bone-image. The fitness values obtained are used to choose potentially superior solutions during the optimization process.

Pseudo-code: FFOA

```

START
1. Initialize population of Fennec foxes randomly
2. Set maximum iterations MaxIter
3. Evaluate fitness of each fox
4. Identify best fox (global best)
5. FOR iter = 1 to MaxIter DO
6.   FOR each fox i in population DO
7.     Generate random coefficients
8.
9.     IF exploration condition is satisfied THEN
10.      Update position using long-distance movement
11.     ELSE
12.      Update position using local exploitation
13.     END IF
14.   Apply boundary constraints
15.   Evaluate new fitness
16.
17.   IF new fitness is better than old fitness THEN
18.     Update fox position
19.   END IF
20. END FOR
21. Update global best fox
22.
23. IF stopping criteria is satisfied THEN
24.   BREAK
25. END IF
26. END FOR
27. Output best solution
28. Output best fitness value
29. Store convergence history
30. END

```

The FFA begins with simulating the fennec fox's digging behavior. This is a local phase of exploration where each solution only explores a small neighbourhood of its current position. This means that it make fine adjustments to the parameters that affect the subtle patterns of osteoporosis. The local update of a decision variable is calculated using the following formula,

$$x_{i,j}^{p1} = x_{i,j} + (2r - 1) \cdot R_{i,j} \quad (17)$$

As the iterations of the algorithm progress, the neighbourhood radius of a solution decreases. This provides a mechanism to explore more broadly in the initial iterations of the FFA, while allowing precise refinement in the later iterations. If the new solution has a better fitness than the current solution and is given by,

$$R_{i,j} = \alpha \cdot \left(1 - \frac{t}{T}\right) \cdot x_{i,j} \quad (18)$$

$$X_i = \begin{cases} x_i^{p1} & \text{if } F_i^{p1} < F_i \\ x_i & \text{otherwise} \end{cases} \quad (19)$$

The current solution is updated with the new solution. The FFA includes an escape mechanism that acts as a global search mechanism. Each solution is encouraged to move towards or away from randomly selected solutions. This simulates the sudden changes in direction a fennec fox employs to evade predators. A random target position is selected as,

$$X_i^{rand} = X_k, k \in \{1, 2, \dots, N\} \quad (20)$$

Global exploration and avoidance of local optima occur in the second phase when the algorithm simulates the fennec fox's escape from predators. Mathematically, it represents the global exploration for the fennec fox algorithm. The equation for this phase is,

$$x_{i,j}^{p2} = \begin{cases} x_{i,j} + r \cdot (x_{i,j}^{rand} - I \cdot x_{i,j}), & \text{if } F_i^{rand} < F_i \\ x_{i,j} + r \cdot (x_{i,j} - x_{i,j}^{rand}), & \text{otherwise} \end{cases} \quad (21)$$

$$X_i = \begin{cases} X_i^{p2}, & \text{if } F_i^{p2} < F_i \\ X_i, & \text{otherwise} \end{cases} \quad (22)$$

Where, I represents a random integer, and r is a random scaling factor. The FFOAs, objective is to explore the feature space and parameter settings of the classification system so that it reaches the global optimum for the best classification of osteoporosis specified in the selected features and classifier parameters. The flowchart of FFOA is shown in Figure 4.

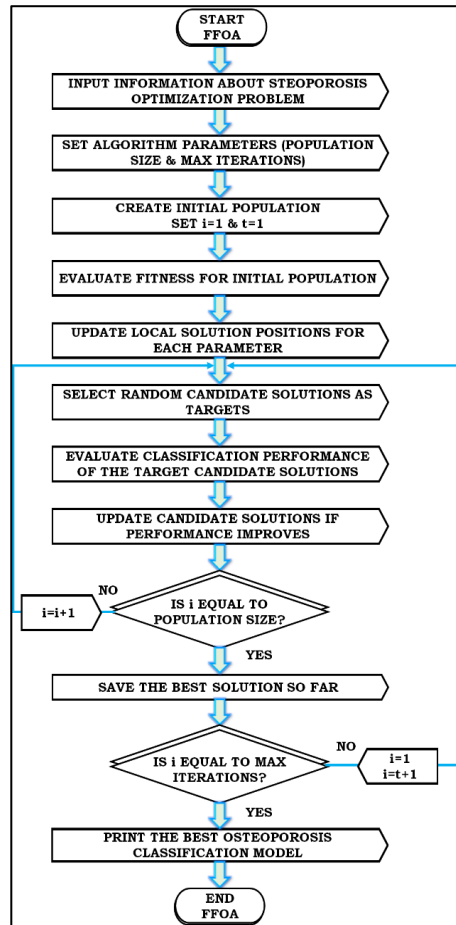


Figure 4: Flowchart of FFOA.
Source: Authors, (2026).

Table 2: Optimization values of FFOA.

SI.NO	PARAMETER NAME	VALUE
1	Population Size	30
2	Maximum Iterations	50
3	Batch Size	32
4	Lower Bound	-10
5	Upper Bound	+10
6	Exploration Probability	0.6
7	Exploitation Probability	0.4
8	Random Movement Factor	[0,1]
9	Fitness Evaluation Function	Objective Function
10	Epoch	80

Source: Authors, (2026).

The FFOA iteratively used for each of these phases allows the active fennec fox algorithm to converge on the optimal classifier for osteoporosis.

IV. RESULT AND DISCUSSION

The aim of this work is the development of a DL system for the automatic classification of osteoporotic changes in the knee, based on knee X-ray images, into three categories such as normal, osteopenia, and osteoporosis, that represent varying degrees of bone density and subject to change due to aging and other conditions. An important feature of this system is the application of a new hybrid model called DualAttention-ViTResNet, it refers to both use of the attention mechanism and vision transformer architecture in conjunction with ResNet Neural Networks to build a superior hybrid model for knee osteoporosis detection based on classification methods. The basis for training this model is the Multi-Class Knee Osteoporosis Dataset, which has an 80:20 split, with 80% of the data being used to train the model and 20% of the data used to evaluate the model's predictive accuracy. The overall intent of implementing the DualAttention-ViTResNet to enhance the ability to identify these osteoporotic changes in knee joints and support medical professionals with accurate diagnosis and ongoing monitoring of bone health is increased accuracy through improved classification.

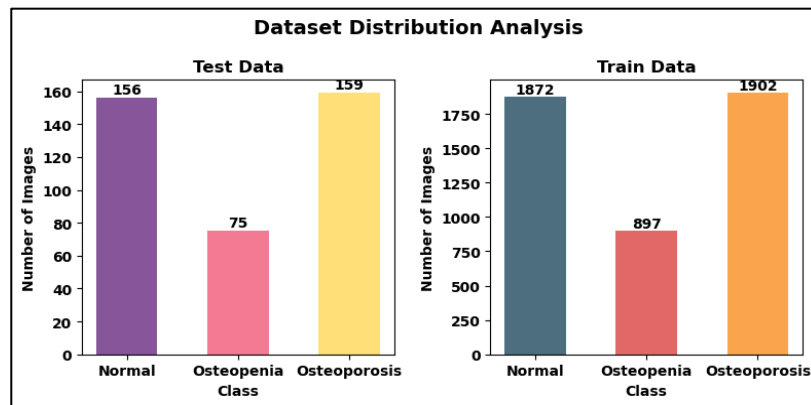


Figure 5: Dataset distribution analysis.

Source: Authors, (2026).

Figure 5 shows the distribution of data analysis, where the test data and train data is mentioned. The dataset revealed a clear imbalance between Normal and Osteoporotic images compared to Osteopenic images in both train and test datasets.

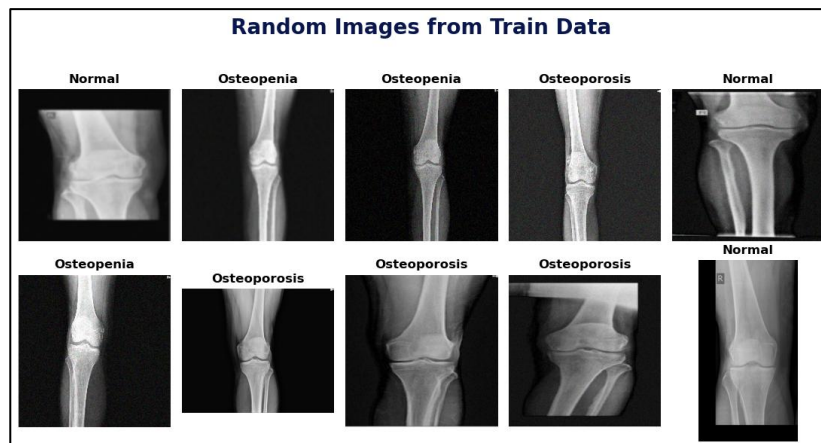


Figure 6: Random images from train data.

Source: Authors, (2026).

The Figure 6 provides examples of Knee X-Rays produced from the training data, comparing normal knees, osteopenic knees, and osteoporotic knees visually illustrating differences in bone density.

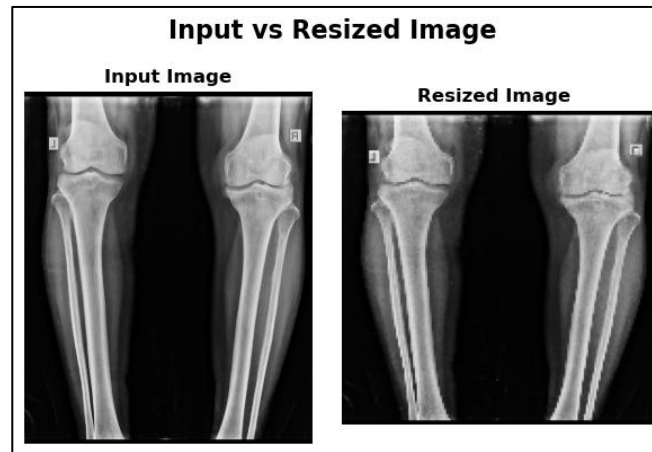


Figure 7: Input image and Resized image.
Source: Authors, (2026).

Figure 7 contains two versions of the same Knee X-Ray, one unaltered and one resized. The initial image appears to be the full-size model used for analysis. The reshaped version is resized in the focus on the actual Knee X-Ray has changed, probable to allow for better review and comparison.

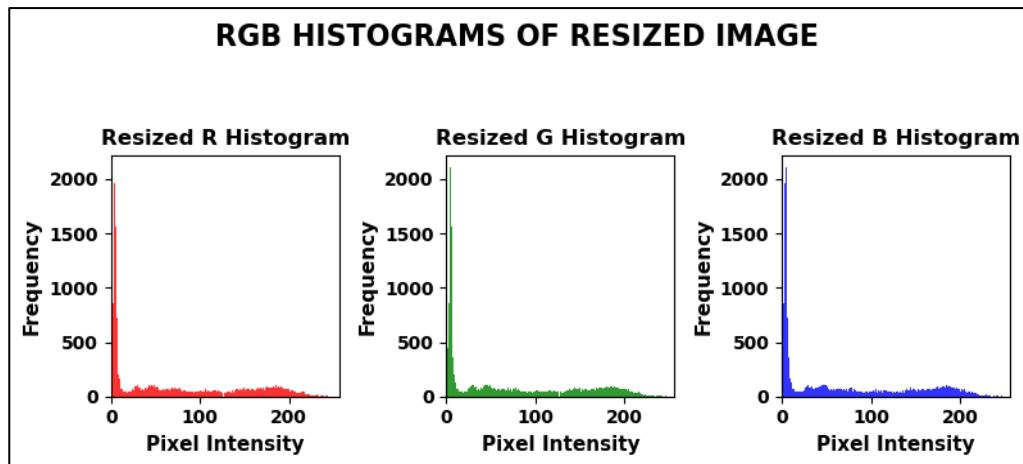


Figure 8: RGB histogram of resized image.
Source: Authors, (2026).

Figure 8 displays RGB histograms related to the resized Knee X-Ray. It is observed from these histograms that there was a distinct clustering of pixel intensity values that were relatively low throughout red, green, and blue channels.

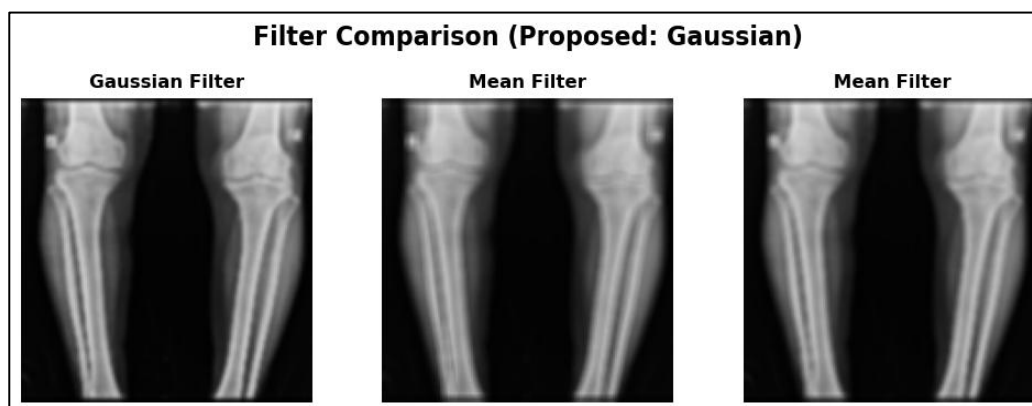


Figure 9: Comparison of filters.
Source: Authors, (2026).

Figure 9 compares the outcomes from the applications of different types of filters to a given X-Ray, Gaussian filter versus mean filter. Although the Gaussian filter produces images that maintain edges with minimal noise, the mean filter creates very smooth results, with significant blurring or lack of definition between the actual knee model and the produced model using a mean filter.

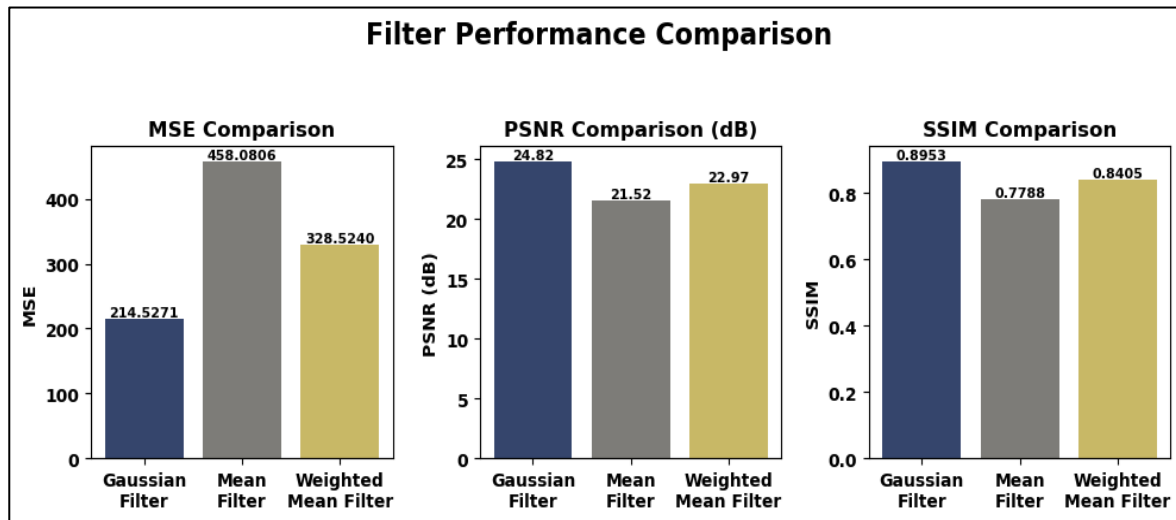


Figure 10: Filter performance comparison.
Source: Authors, (2026).

From Figure 10, it is clearly perceived that the Gaussian filter outperformed all other types of filters through all methods of evaluation, where it produced the lowest MSE, the highest PSNR and the highest SSIM which demonstrates that the quality of the Gaussian filter surpasses that of Mean and Weighted Mean filters.

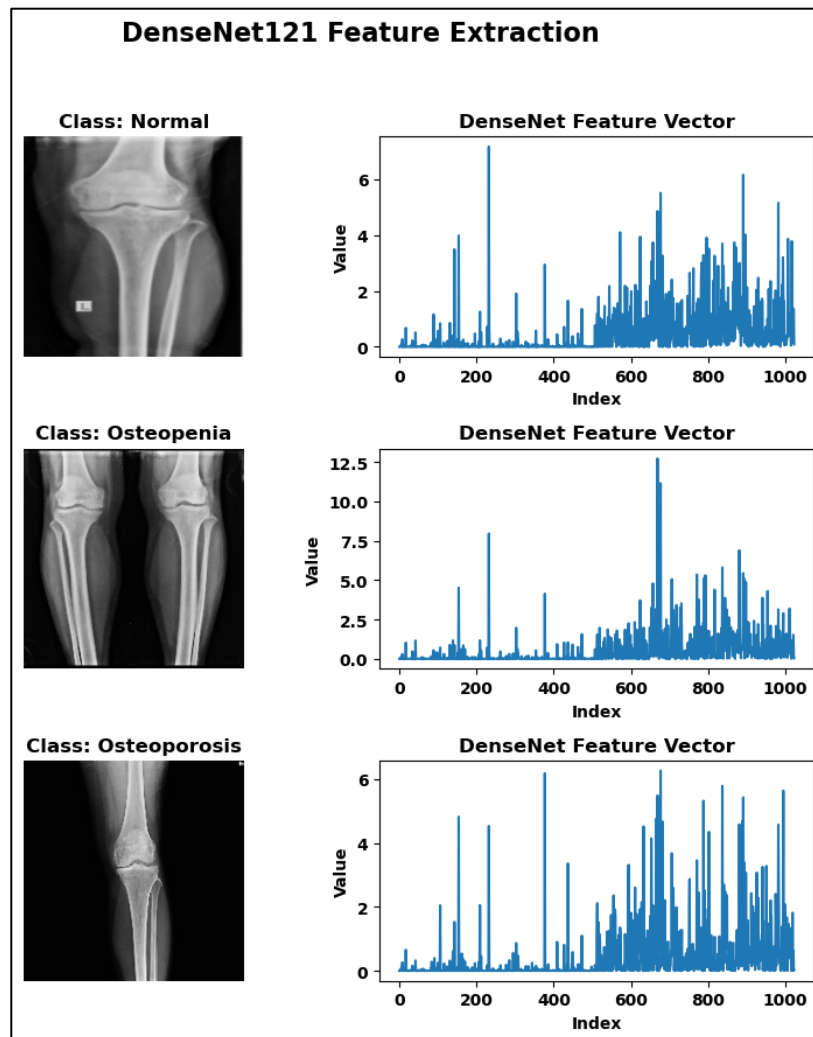


Figure 11: DenseNet-121 Feature extraction.
Source: Authors, (2026).

Figure 11 is the representative of feature extraction using DenseNet121 from three classes of X-ray images such as Normal, Osteopenia and Osteoporosis. It displays the feature vectors associated with the images and exhibit varying patterns of values, which allows for better differentiation between the bone conditions represented by each image.

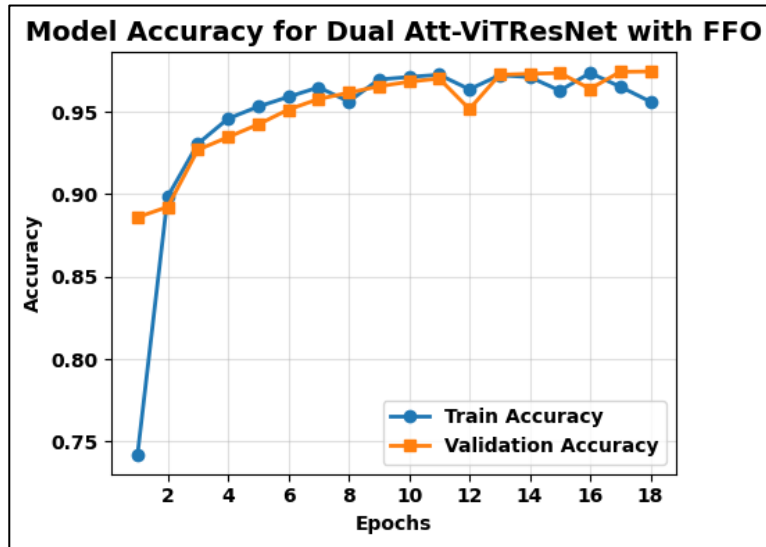


Figure 12: Model accuracy for Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

Figure 12 depicts the accuracy associated with each epoch of training and validation for the Dual Att-ViTResNet with FFO, demonstrating that the accuracy improves quickly during the initial epochs and stabilises towards the end around 96–97% and that the training and validation curves are closely aligned, which indicates a good amount of convergence and minimal overfitting was exhibited.

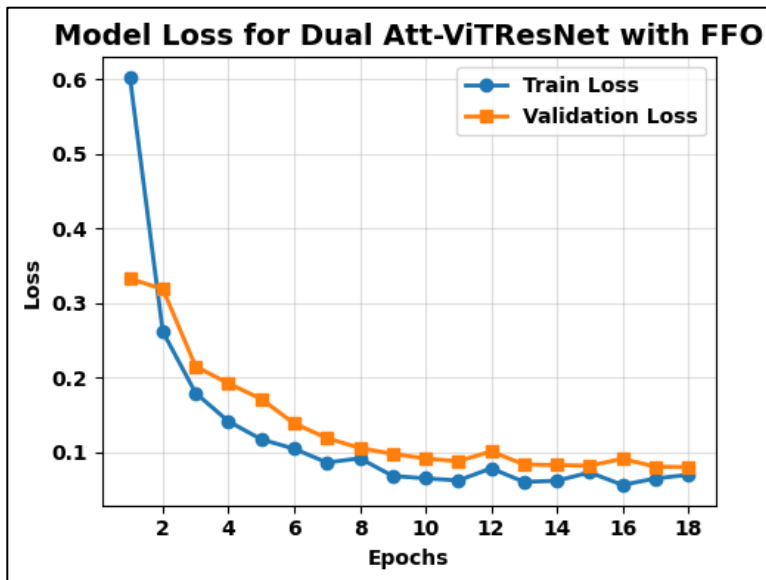


Figure 13: Model loss for Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

Figure 13 represent the decaying loss associated with training and validation over time, with both beginning to stabilise at much lower loss values after a number of epochs have taken place, which demonstrates the model is learning more effectively.

Table 3: Model performance for Dual Att-ViTResNet with FFO.

SI.NO	METRICS	VALUES
1	Accuracy	0.9744
2	Precision	0.9767
3	Recall	0.9789
4	F1-Score	0.9777
5	Specificity	0.9861
6	AUC-ROC	0.9991
7	AUC-PR	0.9987
8	MCC	0.9600
9	Cohen's Kappa	0.9598

Source: Authors, (2026).

Table 3 depicts the model performance of Dual Att-ViTResNet with FFO. The model displays exceptional performance as evidenced by its strong prediction accuracy, balanced precision and recall, superior ability to discriminate between the various classes, and very high correlation between the predictions and reality.

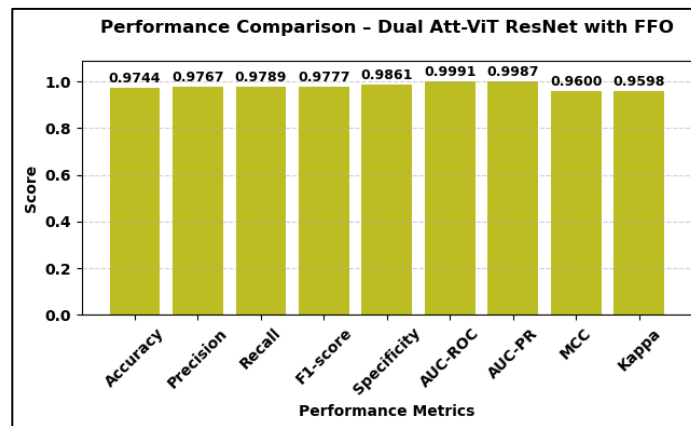


Figure 14: Performance comparison of Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

Figure 14 represents the performance of the Dual Att-Vit ResNet+FFO model relative to all performance metrics. The Dual Att-Vit ResNet+FFO model has a high score for the Accuracy, Precision, Recall, F1 Score, Specificity, AUC-ROC, AUC-PR, MCC and Kappa metrics. These metrics indicate the model's high level of performance across these evaluation dimensions.

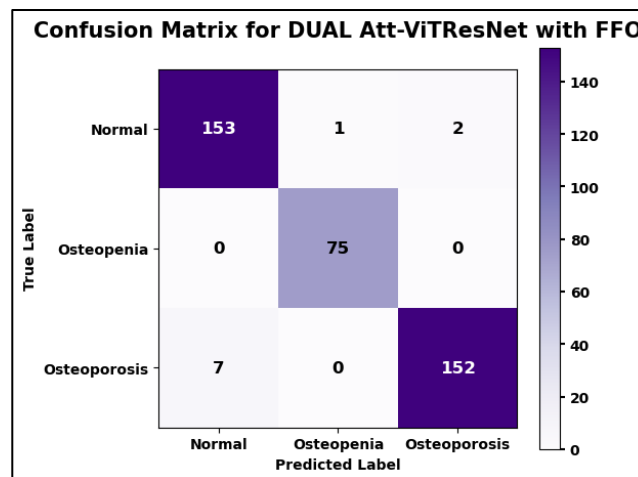


Figure 15: Confusion matrix of Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

The Confusion Matrix of the Dual Att-Vit ResNet + FFO model demonstrates that there is a relatively high number of correctly predicted classifications which is demonstrated in Figure 15. While the Confusion Matrix contain a few incorrectly classified cases, the overall trend indicates that the model accurately classifies individuals into their normal, Osteopenia and Osteoporosis categories.

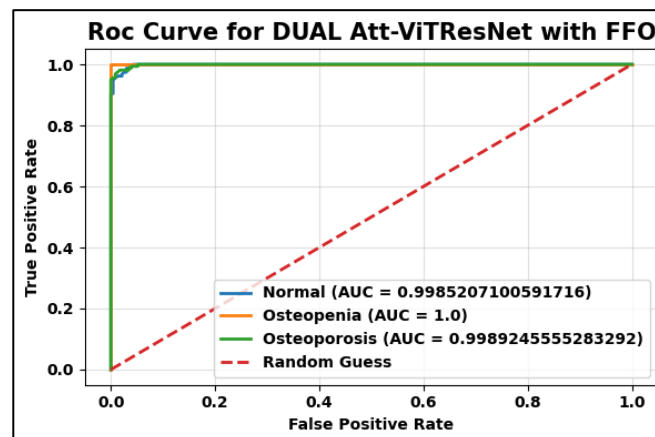


Figure 16: ROC curve for Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

Figure 16 demonstrates a very high level of classification for each class being classified by the Dual Att-ViT + FFO Model. Each of the Normal Osteopenia and Osteoporosis classes has an AUC value of close to 1.0 and therefore demonstrates excellent classification performance compared to random selection.

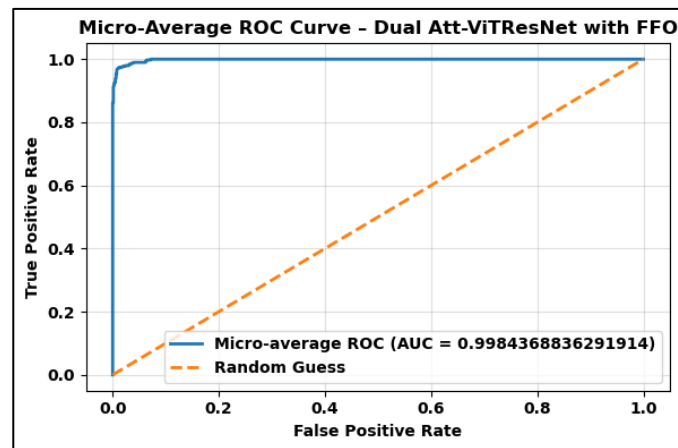


Figure 17: Micro-average ROC curve for Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

Figure 17 demonstrates the Micro-Average ROC Curve of the Dual Att-ViT + FFO Model, with an AUC value of 0.998. This indicates that the model will classify nearly perfectly for micro-averaged data, as this curve is close to the top left corner and well above the baseline for random selection.

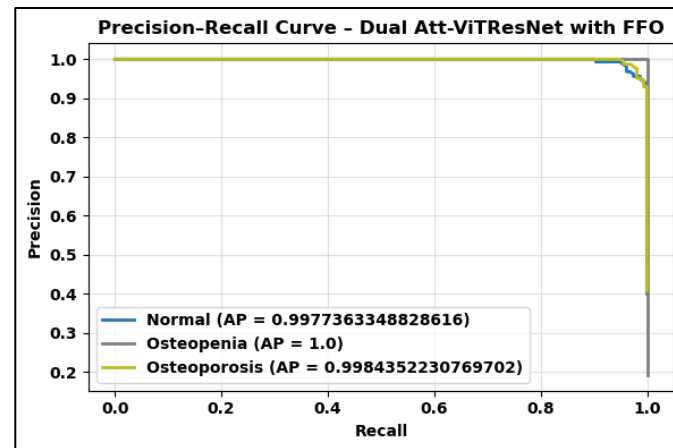


Figure 18: Precision-Recall curve for Dual Att-ViTResNet with FFO.
Source: Authors, (2026).

The Precision-Recall Curves of the Dual Att-ViT + FFO model establish high classification performance for all three classes which depicted in Figure 18. Each class has an average precision value of approximately 0.997 (Normal), 1.0 (Osteopenia) and 0.998 (Osteoporosis). The Precision-Recommendation curves remain close to the value of 1.0 for most levels of Recall and therefore provide a high degree of confidence in the model's classification predictions.

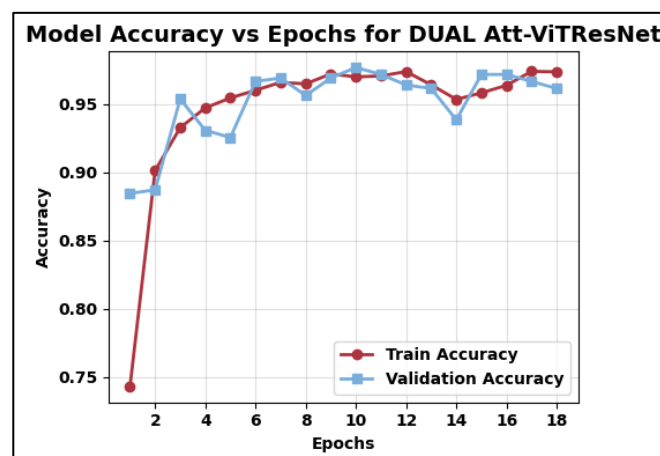


Figure 19: Comparison of model accuracy and epochs for Dual Att-ViTResNet.
Source: Authors, (2026).

Figure 19 shows that the training accuracy and validation accuracy improved rapidly during training accuracy and validation accuracy. Both remained relatively stable across epochs. This indicates that the model is converging well and is not experiencing much overfitting during training.

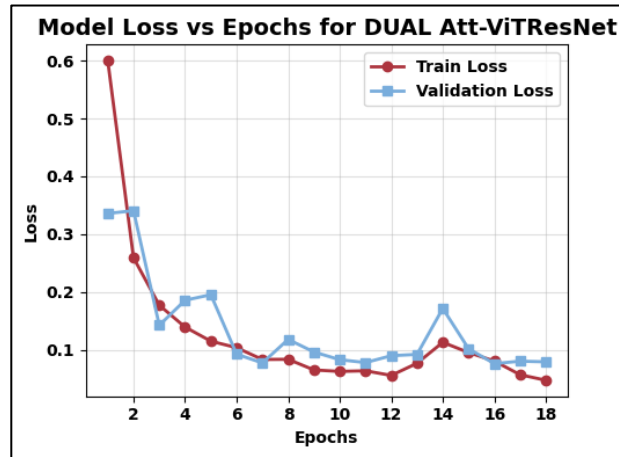


Figure 20: Comparison of model loss and epochs for Dual Att-ViTResNet.

Source: Authors, (2026).

In Figure 20, training loss decreased steadily over the course of training. Though, validation loss started to decrease but started to rise after a few epochs. This indicates the possibility of overfitting occurring after the first few epochs.

Table 4: Model performance for Dual Att-ViTResNet.

SLNO	METRICS	VALUES
1	Accuracy	0.9615
2	Precision	0.9618
3	Recall	0.9682
4	F1-Score	0.9647
5	Specificity	0.9796
6	AUC-ROC	0.9977
7	AUC-PR	0.9965
8	MCC	0.9400
9	Cohen's Kappa	0.9397

Source: Authors, (2026).

Table 4 achieved very strong and consistent results relative to its performance, indicating that the model is very reliable for classifying data accurately with high levels of accuracy, precision, recall, class separation and agreement between model predictions and true labels.

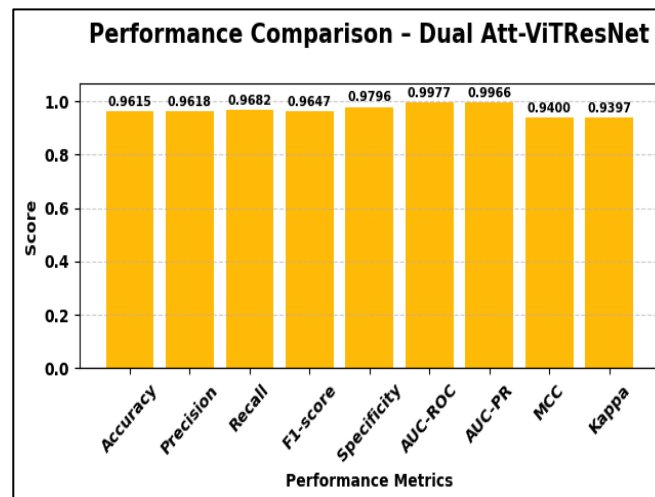


Figure 21: Performance comparison of Dual Att-ViTResNet.

Source: Authors, (2026).

Figure 21 achieved very high scores in all performance metrics as shown in the bar chart, demonstrating very high AUC-ROC and AUC-PR, very high accuracy, precision, recall, F1-score and so on demonstrating a model that is well-balanced and extremely robust.

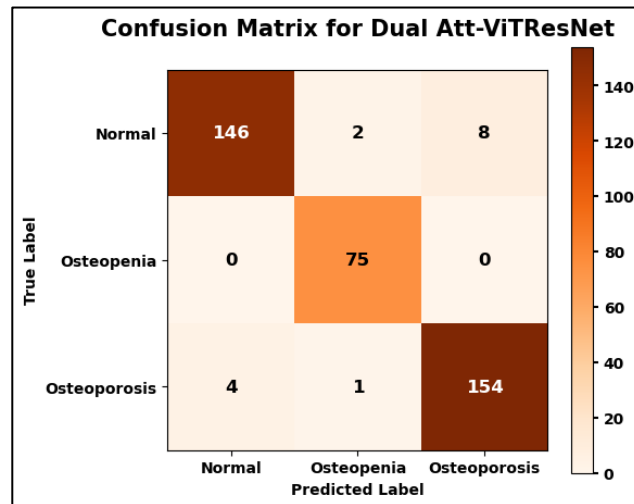


Figure 22: Confusion matrix for Dual Att-ViTResNet.
Source: Authors, (2026).

In Figure 22, the confusion matrix indicates that the model performs well and accurately across the spectrum of classes, but particularly well when classifying people with osteopenia.

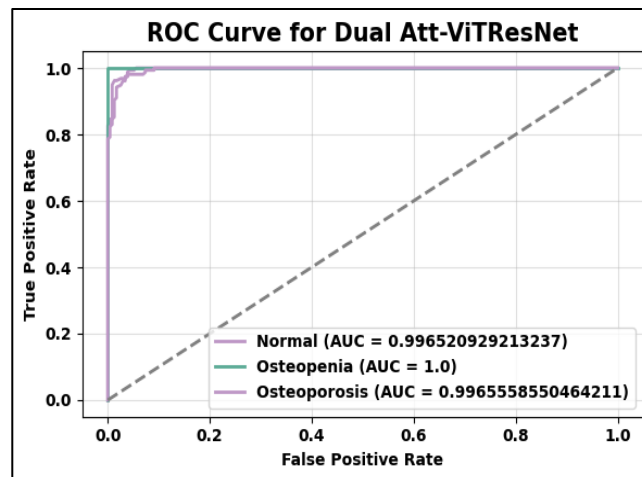


Figure 23: ROC curve for Dual Att-ViTResNet.
Source: Authors, (2026).

The ROC curves in Figure 23 demonstrate excellent performance of Dual Att-ViTResNet in classifying classes of normal, osteopenia and osteoporosis and require AUC values, demonstrating a high degree of confidence and accuracy when making these classifications.

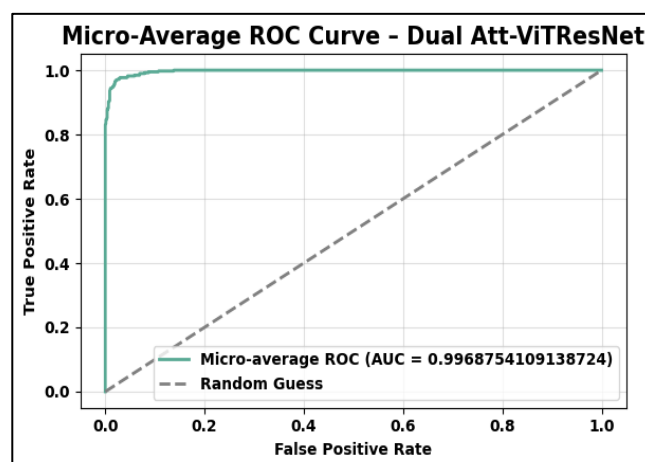


Figure 24: Micro-average ROC curve for Dual Att-ViTResNet.
Source: Authors, (2026).

In Figure 24, the micro-average ROC curve indicates that the Dual Att-ViTResNet has achieved difficulties and excels in the multi-class classification of the low-incidence classes with a higher signal-to-noise ratio than other models.

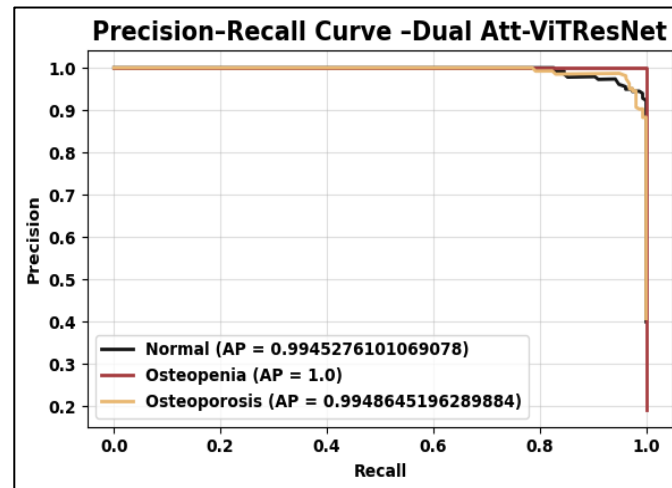


Figure 25: Precision-Recall curve for Dual Att-ViTResNet.

Source: Authors, (2026).

The Dual Att-ViTResNet approach has shown consistent performance over time according to the Precision-Recall Curve in Figure 25, with very high average precision values for each class across nearly the full range of recall. This indicates the Dual Att-ViTResNet has excellent classification capabilities.

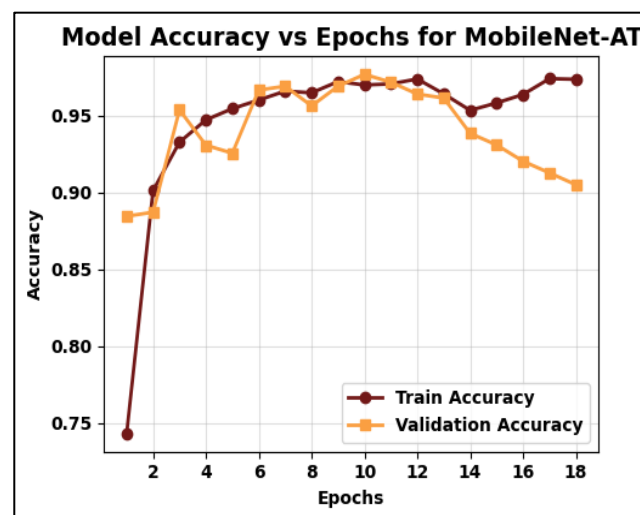


Figure 26: Comparison of model accuracy and epochs for MobileNet-AT.

Source: Authors, (2026).

In Figure 26, the accuracy of MobileNet-AT increases steadily during training, however there is variability in validation accuracy after reaching peak accuracy that suggests there are some overfitting to the training data set.

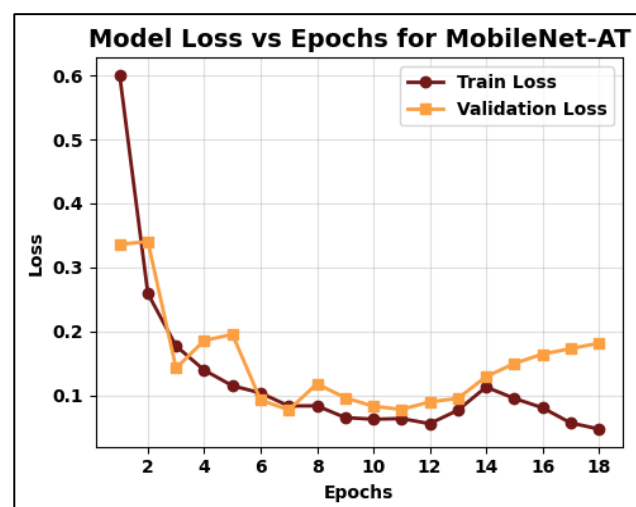


Figure 27: Comparison of model loss and epochs for MobileNet-AT.

Source: Authors, (2026).

In Figure 27, the loss for MobileNet-AT decreases rapidly during the early epochs then stabilizes at about epoch 3. Validation loss decreases also but does level out around epoch 10. The slight difference in the two curves suggests possible overfitting occurring while MobileNet-AT was being trained.

Table 5: Model performance for MobileNet-AT.

SLNO	METRICS	VALUES
1	Accuracy	0.9051
2	Precision	0.8990
3	Recall	0.9217
4	F1-Score	0.9084
5	Specificity	0.9515
6	AUC-ROC	0.9831
7	AUC-PR	0.9772
8	MCC	0.8535
9	Cohen's Kappa	0.8525

Source: Authors, (2026).

According to Table 5, MobileNet-AT shows consistent performance across multiple evaluation metrics including high levels of accuracy and precision, as well as good balance between precision and recall and discrimination ability.

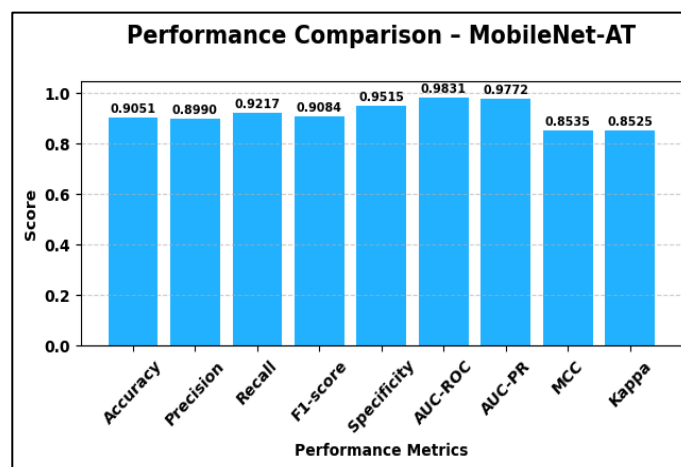


Figure 28: Performance comparison of MobileNet-AT.

Source: Authors, (2026).

Figure 28 displays a consistent high levels of performance across multiple metrics. All metrics show a similar pattern of consistent performance including the highest levels of AUC-ROC and AUC-PR respectively, and specificity indicating high levels of classification capability and discrimination ability.

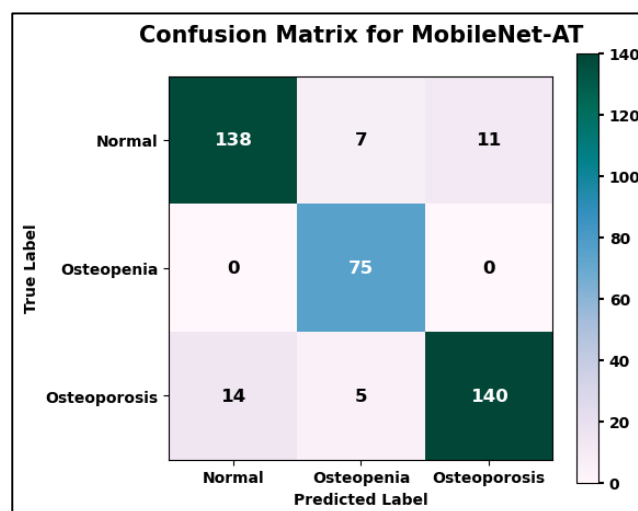


Figure 29: Confusion matrix for MobileNet-AT.

Source: Authors, (2026).

In Figure 29, the confusion matrix for MobileNet-AT indicates that MobileNet-AT accurately classify the vast majority of patients diagnosed as Normal, Osteopenic, and Osteoporotic with few incorrect classifications. This demonstrates strong inclusive performance for the MobileNet-AT model.

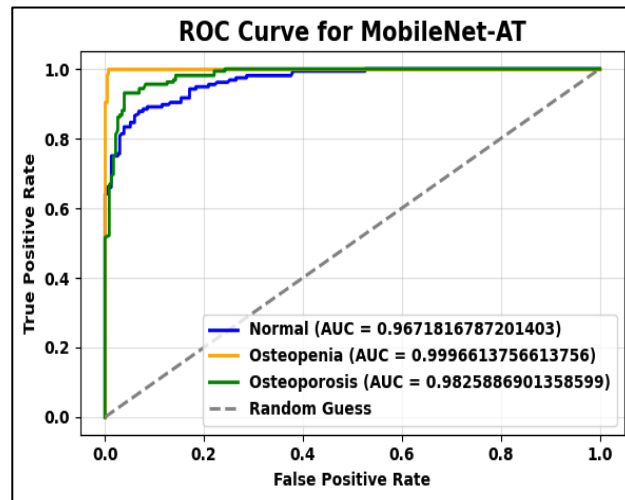


Figure 30: ROC curve for MobileNet-AT.
Source: Authors, (2026).

In Figure 30, the classification ability of MobileNet-AT has been demonstrated using the ROC curves, which show an AUC of approximately 0.967 for Normal, 0.99 for Osteopenia and approximately 0.98 for Osteoporosis. Thus, the MobileNet-AT Model is able to differentiate well between the different classes of subjects.

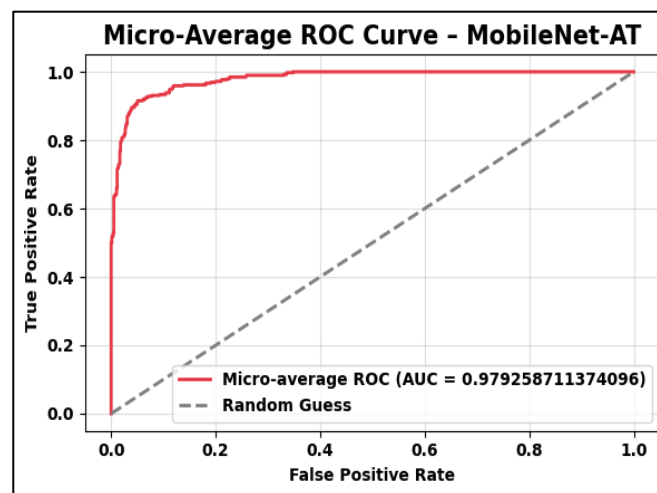


Figure 31: Micro- Average ROC curve for MobileNet-AT.
Source: Authors, (2026).

In Figure 31, the Micro-Average ROC of the MobileNet-AT model shows overall classification performance to be excellent with an AUC of about 0.979, thus demonstrating high classification accuracy for all classes of subjects.

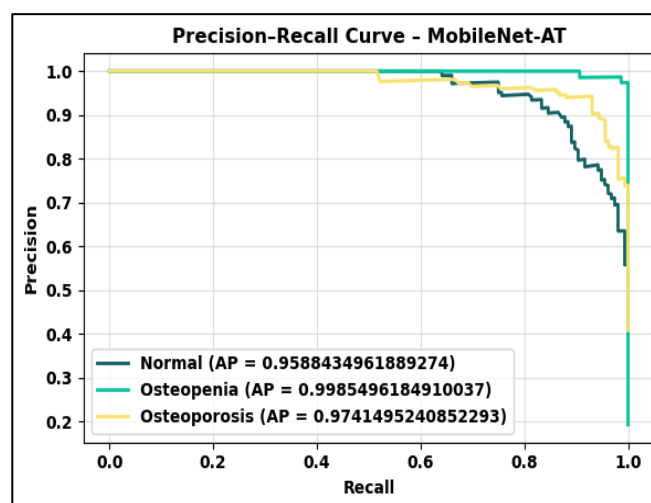


Figure 32: Precision-Recall curve for MobileNet-AT.
Source: Authors, (2026).

In Figure 32, the Precision-Recall curves of the MobileNet-AT model for Normal, Osteopenia and Osteoporosis provide the following rankings, Osteopenia provides the best performance ($AP = 0.998$), followed by Normal, and then Osteoporosis.

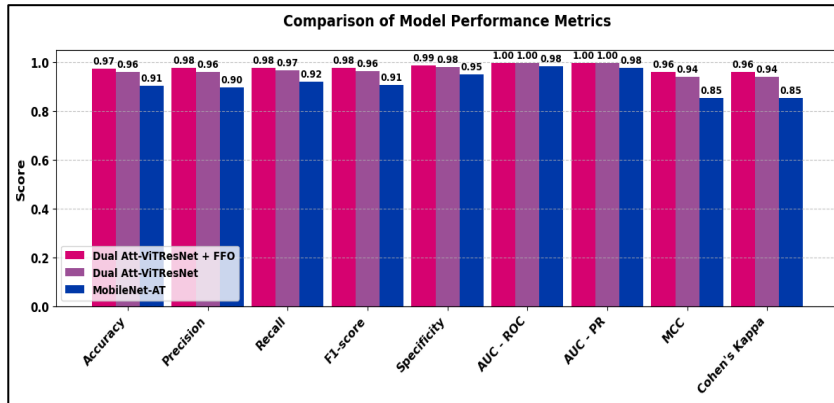


Figure 33: Comparison of model performance metrics.

Source: Authors, (2026).

Figure 33 compares the three models across three metrics. Dual Att-ViTResNet + FFO is shown to consistently outperform both Dual Att-ViTResNet and MobileNet-AT.

Table 6: Comparison of accuracy for various methods.

SI. NO	METHODS	ACCURACY
1	OsteoNet [23]	0.741
2	ViT-16 [21]	0.880
3	MobileNet [24]	0.940
4	ResNet+ GRU [25]	0.956
5	PROPOSED	0.974

Source: Authors, (2026).

Table 6 shows that the proposed method outperformed OsteoNet, ViT-16, MobileNet, and ResNet + GRU to achieve an accuracy of 0.974.

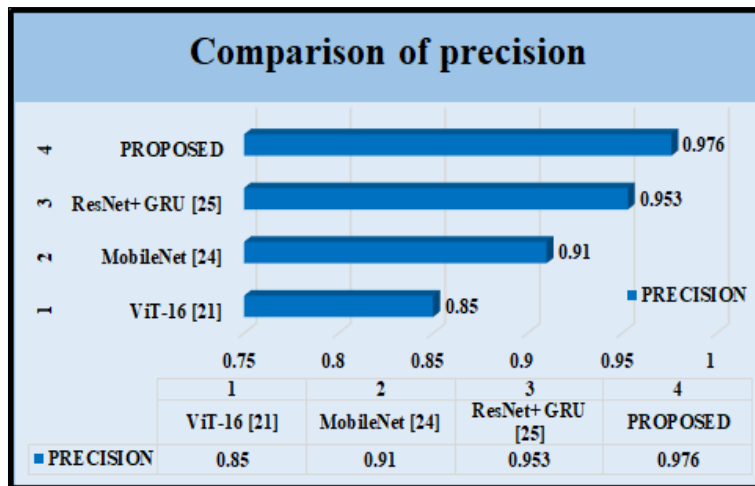


Figure 34: Comparative analysis of precision.

Source: Authors, (2026).

The precision results are shown in Figure 34. The proposed method outperformed all other methods with a precision of 0.976, followed by ResNet + GRU, MobileNet, and ViT-16, respectively.

Table 7: Comparison of Recall and F1-Score for various methods.

SI.NO	METHODS	RECALL	F1-SCORE
1	ViT-16 [21]	0.880	0.860
2	ResNet+ GRU [25]	0.959	0.954
3	PROPOSED	0.978	0.977

Source: Authors, (2026).

According to the Table 7, the suggested algorithm yield the greatest recall (0.978) and F1-score (0.977). The results achieved by the suggested method outperform those of ViT-16 and ResNet+GRU.

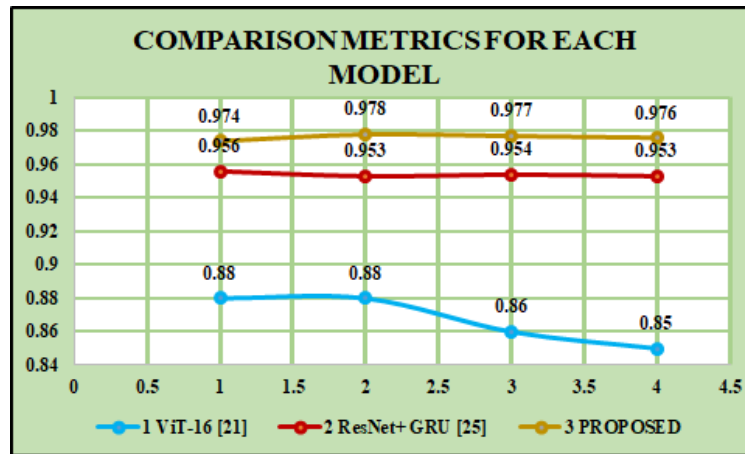


Figure 35: Comparison metrics for each model.
Source: Authors, (2026).

In Figure 35, three models are compared in terms of their performance, the ViT-16, ResNet+GRU, and the Proposed Model. The proposed model has consistently produced superior metrics when compared to the ViT-16 and ResNet+GRU, which have lower and decreasing performance metrics over time.

V. CONCLUSION

The automated classification of osteoporosis was achieved by developing a process based on the proposed DualAtt-ViTResNet framework using the FFOA. The suggested system is integrated by a number of different processes, all employed together to enhance the quality of the classification at each step. The process of preprocessing images, through Resizing and Gaussian Filtering improves the quality of the input images along with reducing noise, which leads to consistently being able to create a properly formatted input for the rest of the system. The addition of the StyleGAN-based data augmentation creates an enriched dataset that includes different classes of images, allowing for better generalization from the training data while minimizing the risk of overfitting. Extracting deep, meaningful features from the input images is primarily performed by the DenseNet-121, which successfully extracts important characteristics that describe the images. The DualAtt-ViTResNet takes advantage of those features and captures both the local detail and the broader context of the image, allowing for more accurate processing and classification of images. Finally, the hyperparameter optimization process uses the FFOA to optimize the learning rate and enhance model accuracy, making the entire system extremely robust and reliable. Overall, by integrating all the stages, the model yields a higher accuracy of 0.9744, precision of 0.9767, Recall of 0.9789 and F1-Score of 0.9777, thus providing reliable, effective, and clinically useful automated classification of osteoporosis.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: V. Venkatachalam, N. Subash, P. Archana Devi, Saranya S, R. Tharwin Kumar, Chokkan Priyadharshini.

Methodology: N. Subash, P. Archana Devi, Saranya S, R. Tharwin Kumar, Chokkan Priyadharshini.

Investigation: V. Venkatachalam, N. Subash, P. Archana Devi.

Discussion of results: V. Venkatachalam, N. Subash, P. Archana Devi, Saranya S, R. Tharwin Kumar, Chokkan Priyadharshini.

Writing – Original Draft: V. Venkatachalam.

Writing – Review and Editing: N. Subash, P. Archana Devi, Saranya S, R. Tharwin Kumar, Chokkan Priyadharshini.

Resources: N. Subash, P. Archana Devi, Saranya S.

Supervision: N. Subash, P. Archana Devi, Saranya S, R. Tharwin Kumar, Chokkan Priyadharshini.

Approval of the final text: V. Venkatachalam.

VII. REFERENCES

- [1] S. Wang, X. Tong, Q. Cheng, Q. Xiao, J. Cui, J. Li, Y. Liu, and X. Fang, "Fully automated deep learning system for osteoporosis screening using chest computed tomography images," *Quantitative Imaging in Medicine and Surgery*, vol.14, no. 4, pp. 2816, 2024.
- [2] V. Kalavakonda, S. Mohamed, L. Abhay, and S. Muthu, "Automated screening of hip X-rays for osteoporosis by Singh's index using machine learning algorithms," *Indian Journal of Orthopaedics*, vol. 58, no. 10, pp. 1449-1457, 2024.
- [3] G. Gatineau, E. Shevroja, C. Vendrami, E. Gonzalez-Rodriguez, W. D. Leslie, O. Lamy, and D. Hans, "Development and reporting of artificial intelligence in osteoporosis management," *Journal of Bone and Mineral Research*, vol. 39, no. 11, pp. 1553-1573, 2024.
- [4] C. Du, J. He, Q. Cheng, M. Hu, J. Zhang, J. Shen, S. Wang, Y. Liu, J. Li, and W. Wei, "Automated opportunistic screening for osteoporosis using deep learning-based automatic segmentation and radiomics on proximal femur images from low-dose abdominal CT," *BMC Musculoskeletal Disorders*, vol. 26, no. 1, pp. 378, 2025.
- [5] R. A. Karthika, A. Rohini, P. Muthukumar, P. Mahmud, and T. S. Babu, "A Resource Efficient Machine Learning Pipeline to Detect Atrial Fibrillation," *2025 8th International Conference on Circuit, Power & Computing Technologies (ICCPCT)*, Kollam, India, pp. 95-100, 2025.
- [6] M. Y. Ichahane, and N. Assad, "Multi-Model Attention-Enhanced CNN Ensemble for Automated Osteoporosis Detection in Radiographic Knee Images," *International Journal of Intelligent Engineering & Systems*, vol. 18, no. 3, 2025.

- [7] U. Hwang, C. H. Lee, and K. Yoon, "Osteoporosis prediction from hand X-ray images using segmentation-for-classification and self-supervised learning," *Scientific Reports*, vol. 15, no. 1, pp. 32964, 2025.
- [8] S. Sharma, K. V. M. Krishna, S. S. Joshi, M. Radhakrishnan, S. Palaniappan, S. Dussa, and R. Banerjee, "Laser based additive manufacturing of tungsten: Multi-scale thermo-kinetic and thermo-mechanical computational model and experiments," *Acta Materialia*, vol. 259, pp. 119244, 2023.
- [9] H. O. El-Sisi, F.E. El-Gamal, and N. A. Hikal, "A non-invasive computer-aided personalized diagnosis system for Osteopenia and Osteoporosis," *Egyptian Informatics Journal*, vol. 29, pp. 100634, 2025.
- [10] A. M. Sarhan, M. Gobara, S. Yasser, Z. Elsayed, G. Sherif, N. Moataz, Y. Yasir, and E. Moustafa, "Knee osteoporosis diagnosis based on deep learning," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, pp. 241, 2024.
- [11] A. Siddiqua, R. Hasan, A. Rahman, and A. S. M. Miah, "Computer-Aided Osteoporosis Diagnosis Using Transfer Learning with Enhanced Features from Stacked Deep Learning Modules," *arXiv preprint arXiv: 2412.09330*, 2024.
- [12] W. Liawrungrueang, S. T. Cho, V. Kotheeranurak, K. Jitpakdee, P. Kim, and P. Sarasombath, "Osteoporotic vertebral compression fracture (OVCF) detection using artificial neural networks model based on the AO spine-DGOU osteoporotic fracture classification system," *North American Spine Society Journal (NASSJ)*, vol. 19, pp. 100515, 2024.
- [13] K. Karthika, and R. Devi Priya, "Image processing-based protection of privacy data in cloud using NTRU algorithm," *Signal, Image and Video Processing*, vol. 18, no. 5, 2024.
- [14] S. M. Baik, H. J. Kwon, Y. Kim, J. Lee, Y. H. Park, and D. J. Park, "Machine learning model for osteoporosis diagnosis based on bone turnover markers," *Health Informatics Journal*, vol. 30, no. 3, 14604582241270778, 2024.
- [15] D. J. Tsai, C. Lin, C. S. Lin, C. C. Lee, C. H. Wang, and W. H. Fang, "Artificial intelligence-enabled chest X-ray classifies osteoporosis and identifies mortality risk," *Journal of medical systems*, vol. 48, no. 1, pp. 12, 2024.
- [16] A. J. Shawon, I. I. M. Gazi, H. R. Hiya, and A. Roy, "Osteoporosis Prediction Using VGG16 and ResNet50," *International Journal of Innovative Science and Research Technology (IJISRT) IJISRT24APR2565*, pp. 2489-2492, 2024.
- [17] N. M. Cross, J. Perry, Q. Dong, G. Luo, J. Renslo, B. C. Chang, N. E. Lane, L. Marshall, S. K. Johnston, D. R. Haynor, and J. G. Jarvik, "Subject-level spinal osteoporotic fracture prediction combining deep learning vertebral outputs and limited demographic data," *Archives of Osteoporosis*, vol. 19, no. 1, pp. 87, 2024.
- [18] X. Tong, S. Wang, J. Zhang, Y. Fan, Y. Liu, and W. Wei, "Automatic osteoporosis screening system using radiomics and deep learning from low-dose chest CT images," *Bioengineering*, vol. 11, no. 1, pp. 50, 2024.
- [19] V. Madhivanan, K. Vijaya, A. Lal, S. Rithika, S. K. Subramaniam, and M. Sameer, "Unsupervised Machine Learning for Osteoporosis Diagnosis Using Singh Index Clustering on Hip Radiographs," *arXiv preprint arXiv: 2411.15253*, 2024.
- [20] A. K. Gebre, M. Sim, S. Z. Gilani, A. Saleem, C. Smith, D. Hans, S. Reid, B. A. Monchka, and D. Kimelman, "Automated abdominal aortic calcification and trabecular bone score independently predict incident fracture during routine osteoporosis screening," *Journal of Bone and Mineral Research*, pp. zjaf144, 2025.
- [21] M. U. Islam, A. Fathima, and D. Ghosh, "An Improved Technique for Bone Fracture Classification Using Visual Transformers and Convolutional Neural Networks," Available at SSRN 5191555, 2025.
- [22] M. J. A. Rasool, S. Ahmad, U. Sabina, and T. K. Whangbo, "Konet: Toward a weighted ensemble learning model for knee osteoporosis classification," *IEEE Access*, vol. 12, pp. 5731-5742, 2024.
- [23] A. W. Muzaffar, F. Riaz, and M. Tahir, "OsteoNet-A Framework for identifying Osteoporosis in Bone Radiograph Images Using Attention Based VGG Network," 2025.
- [24] S. Thota, P. Kandukuru, M. Sundaram, A. Ali, S. M. Basha, and N. H. Bindu, "Deep Learning based Bone Fracture Detection," In *2024 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*, pp. 1-7. IEEE, 2024.
- [25] A. OM and R. Gunasundari, "Advanced Osteoporosis Prediction from Knee X-rays via Residual Convolution and Recurrent Networks," *International Journal of Intelligent Engineering & Systems*, vol. 18, no. 6, 2025.