International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(9s) 2026 pp. 436-452
https://doi.org/10.70917/ijcisim-2026-3450
Article

# Swarm-Optimized Multi-Scale Deep Learning for Accurate Knee Osteoarthritis Assessment

**Sujeet More[1,*], Ravi Kalkundri[2], Geetika Narang[3], Amol Bhosale[4], Sneha Tirth5, Rupali Maske6**

[1,*] Department of Computer Engineering, Trinity College of Engineering & Research, Pune, India sujeetmore7@gmail.com
[2,*] Department of Computer Science and Engineering, Gogte Institute of Technology, Belagavi, India kalkundri.ravi05@gmail.com
[3]Department of Computer Engineering, Trinity College of Engineering & Research, Pune, India geetikanaang.tcoer@kjei.edu.in
[4]Department of Electronics and Telecommunication, Trinity College of Engineering & Research, Pune, India amolbhosale.tcoer@kjei.edu.in
5Department of Computer Engineering, Trinity College of Engineering & Research, Pune, India snehatirth.tcoer@kjei.edu.in
6Department of Computer Engineering, Trinity College of Engineering & Research, Pune, India rupalimaske.tcoer@kjei.edu.in
*Corresponding author: sujeetmore7@gmail.com, kalkundri.ravi05@gmail.com

**Abstract:** Knee Osteoarthritis is a weakening joint disorder, which mostly destroys the knee articular cartilage and in its cruel phases causes severe pain and frequently mains to complete joint replacement. It mainly affects aging, overweight, and those who survive in a sedentary lifestyle. Prompt analysis is vital for pathology and medical treatments. Professional radiologists visually inspect the joint of knee areas and discover the structures utilizing image processing and computer vision (CV)-based techniques, and might even utilize CAD-based extents for evaluating the severity of knee OA. However, manual approaches are highly experience-based and subjective, and they are difficult and tedious when a huge amount of issues want to be inspected. So, the machine learning and deep learning- based approaches are getting attention and distributed higher accuracy and adequate precision for the recognition and identification of KOA at an initial phase. In this study, we offer an Enhanced Multi-Scale Feature Fusion for Knee Osteoarthritis Severity Detection Using the Dwarf Mongoose Optimization (EMSFF-KOSDDMO) model. The main purpose of EMSFF-KOSDDMO model is to enhance the recognition of knee osteoarthritis severity using medical image data. Initially, the EMSFF-KOSDDMO technique applies non-local means (NLM) filtering for noise removal and adaptive gamma correction with weighting distribution (AGCWD) for contrast enhancement to enhance the knee X-ray image quality. Besides, the presented EMSFF-KOSDDMO model designs a multi-scale feature fusion of EfficientNetB0 with MBConv4 and EfficientNetV2-B0 with MBConv6 for the extraction of feature process to capture rich, multi-scale features with efficient computational performance. For the detection process, the TabNet method has been exploited. At last, the DMOA alters the hyperparameter values of TabNet system optimally and outcomes in greater performance of classification. A wide sort of experimentations are conducted to validate the performance of EMSFF-KOSDDMO system under various evaluation measures. The simulation results suggested that the EMSFF-KOSDDMO system emphasized improvement over other existing models.

**Keywords:** Knee Osteoarthritis Severity Detection; Dwarf Mongoose Optimization; Multi-Scale Feature Fusion; TabNet; Image Preprocessing

## 1. Introduction

Investigation into osteoarthritis (OA) pathology is essential for a lifetime because of its higher economic effect, pain, disability, and serious effect on the patient's existence [1]. OA surpasses physiological and anatomical changes, for cellular stress and deterioration of an extracellular cartilage matrix originate with macro- and micro-injuries. Nevertheless, there are other risk aspects such as absence of exercise, obesity, bone density, gender, genetic predisposition, trauma, and occupational injury [2]. Knee osteoarthritis (KOA) is an illness, which is most prevalent in the aging population and consequence of wearing the articular cartilage from the middle of knee joints [3]. In the medical sector, practitioners utilize the Kellgren-Lawrence (KL) grading method as normal to categorize KOA seriousness from radiographs. Radiographs remain a widely utilized imaging modality due to their affordability and accessibility, even with the advent of more advanced clinical imaging technologies. The seriousness of identification precision is strongly dependent on the experience and carefulness of the specialist

[4]. It believes that the lower reliability of the physician's grading is because of these very minute changes among radiographs of nearby grades. Additionally, it is considered that each specialist might have a diverse suggestion on the radiograph severity grade depending on their understanding and experience [5]. KOA is also an illness that is extremely difficult to identify in the initial stages when the grade difference between zero and one is very minimal.

Nevertheless, early treatment and detection can assist the aging slow down OA progression, and enhance their living standards [6]. There are multiple imaging modalities like Optical Coherence Tomography, MRI, radiography (X-ray), and ultrasound for increased OA analysis have been conventionally chosen, and remain the major attainable devices and golden standard for preliminary knee OA identification [7]. The analytic precision is extremely dependent on the physician's carefulness and experience. Furthermore, the KL grading criteria are very ambiguous. More precise, reliable analysis of the OA phase will ensure that inspective treatments are calculated in patients and the severity range designed by the examiners [8].

In medical practice, holding a steady, automated mechanism to evaluate consecutive radiographic inspections of individual patients allows better tracking of their illness formation [9]. Preceding analysis has advanced automated radiographic classifiers, extending from Machine Learning (ML) methodologies that depend on image information distilled utilizing subject areas from Computer Vision (CV) specialists to Deep Learning (DL) models. DL techniques have accomplished better, but preceding methods have depended on explanations of the joint space [10]. DL is a subdivision of ML but an approaches turn nearby building layered methods to permit a computer to independently implement particular tasks like object detection and classification.

This manuscript designs and develops an Enhanced Multi-Scale Feature Fusion for Knee Osteoarthritis Severity Detection Using Dwarf Mongoose Optimization (EMSFF-KOSDDMO) model. Initially, the EMSFF-KOSDDMO technique applies non-local means (NLM) filtering for noise elimination and AGCWD for contrast enhancement. Besides, the projected EMSFF-KOSDDMO system designs a multi-scale feature fusion of EfficientNetB0 with MBConv4 and EfficientNetV2-B0 with MBConv6 for the extraction of the feature procedure. For the recognition process, the TabNet model has been exploited. At last, the DMOA alters the hyperparameter values of TabNet system and outcomes in greater performance of classification. A wide sort of experimentations are conducted to validate the performance of EMSFF-KOSDDMO system under various evaluation measures.

## 2. Literature Review

Mehdi et al. [11] projected a new DL method that associates texture aspects and shapes to score knee OA seriousness from X-ray imageries. Dual DL techniques are integrated: a Siamese CNN for removing shape aspects from the clinical and lateral section of knee joint and a Discriminant Regularized AE (DRAE) for removing a texture aspects. The eliminated attributes are then fused utilizing the Fully Connected (FC) layer and Compact Bilinear Pooling (CBP) for knee OA recognition depending on the score of KL. The authors [12] utilize a 12-layer CNN to attain DL capability. Data collected from the OAI is utilized to categorize KOA. An advanced CNN structure generated particularly for dual classifications and KOA seriousness employing DL models is the major element of this study. Kadu and Pawar [13] developed an advanced method exploiting the single ability of bilinear CNN (BiCNN). To incorporate bilinear communications in the CNN structure, this technique focuses on taking channel-wise dependency and convoluted spatial in knee radiographic images, consequently enhancing the ability to modify the osteoarthritis progression, mainly in the joint area.

The authors [14] introduce a DL-based structure such as OsteoHRNet that mechanically considers the Knee OA seriousness regarding KL grade classification from X-ray. Additionally, an attention mechanism is integrated to remove the ineffective aspects and enhance further implementation. Fang et al. [15] developed a multi-stage method for identifying meniscus injuries over image segmentations into lesion regions, succeeded by an integrated classification model. The existing work applies a model by which the image sound is reduced, and after that the performance of an upgraded version of U-Net model to implement image segmentation and recognize ROI for possible injury recognition. The feature extraction was attained over the usage of the contour line model. Sivakumar et al. [16] developed a CNN for knee joint identification, providing effective alternate expert-dependent models. This work projects an innovative DL-based CNN approach for forecasting OA development utilizing knee MRI images. This technique also leverages inclusive methods, incorporating data segmentation, pre-processing, feature extraction and collection utilizing GLCM, and classification employing CNN.

Segura et al. [17] examine the usage of CV and AI for an automated classification and detection of KOA utilizing the IKDC classification method. A public database comprising radiographic knee images was employed. Images were handled exploiting LandingLens software, a progressive CV framework enabling AI techniques implementation and development. An ML technique depends on the ConvNext structure CNN was proficient and assessed. Karegoudra et al. [18] comprise an in-depth exploration and inspection of DL applications for knee osteoarthritis recognition and forecasting severity. During this paper, EfficientNet-B6 was selected for its superior performance.

## 3. The Proposed Method

In this manuscript, we offer an EMSFF-KOSDDMO model. The main purpose of EMSFF-KOSDDMO technique is to enhance the recognition of knee osteoarthritis severity using medical image data. It encompasses distinct various phases involved as image pre-processing, multimodal feature extraction, classification using TabNet, and hyperparameter model. Fig. 1 represents the entire flow of EMSFF-KOSDDMO approach.

### *3.1. Image Preprocessing*

Initially, the EMSFF-KOSDDMO technique applies NLM filtering for noise removal and AGCWD for contrast enhancement to enhance the knee X-ray image quality.

#### 3.1.1. Noise Removal

NLM is an effective method for image denoising, which uses the termination existing in original images [19]. It works by average related image patches to successfully subdue noise by maintaining the detail of the images. In this model, we use NLM as a postprocessing stage to improve the denoised images. In detail, for all pixels in the denoised images, they calculate weighting averages of related patches from the complete images, whereas the weightings are established according to the similarities among patches. By combining NLM into the pipeline, we target to reach greater denoised outcomes with improved protection of good image architectures.

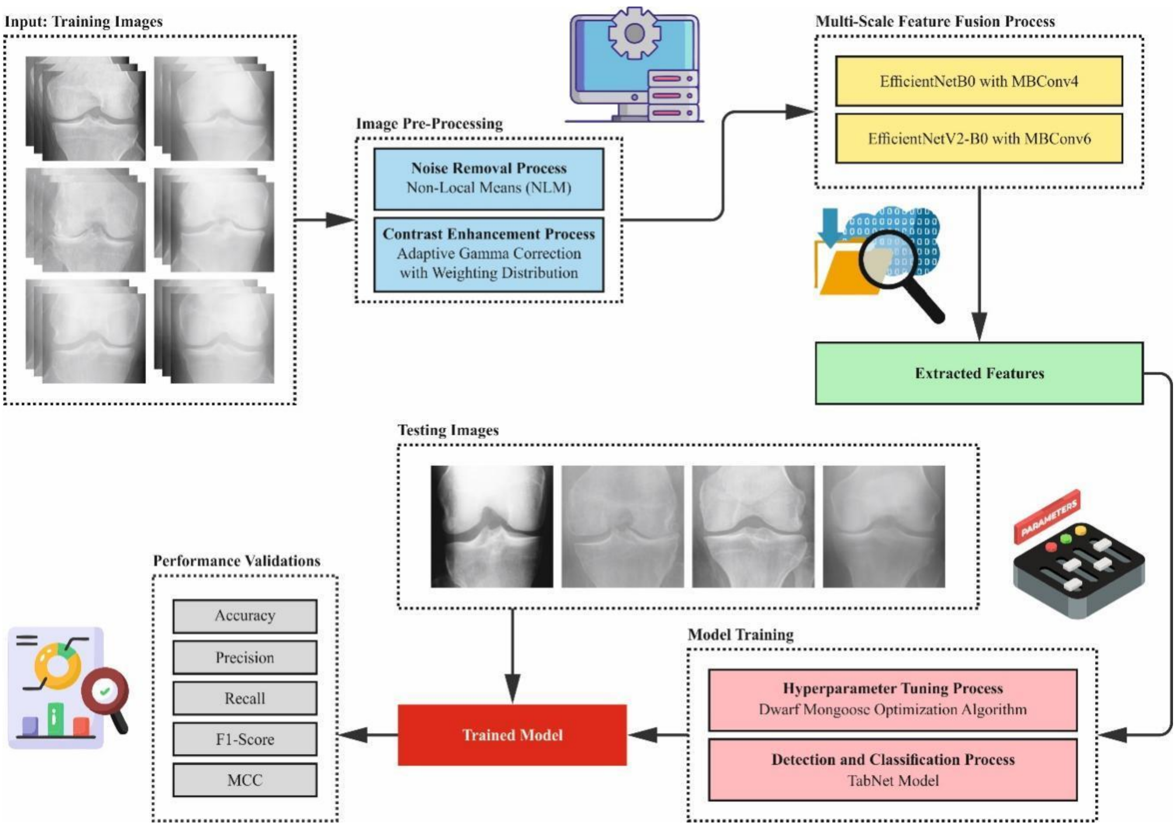


**Fig. 1.** Overall flow of EMSFF-KOSDDMO approach

Provided noisy images $v$, the denoised value $u(i)$ at $ith$ pixels was calculated as weight averages of each pixel $j$ inside an images. The weight allocated to all pixels $j$ relies on similarities amongst pixel's neighborhoods $i$ and $j$. The NLM model is stated as shown:

$$u(i) = \frac{1}{C(i)} \sum_{j \in \Omega} w(i,j) v(j) \tag{1}$$

Whereas $u(i)$ denotes the denoised value of $ith$ pixels , $\Omega$ refers to the collection of each pixel in an image, and $w(ij)$ represents weighted that calculates the similarities amongst the pixel's neighborhoods $i$ and $j$. $C(i)$ signifies the normalization feature, and $v(j)$ stands for noise image value at $jth$ pixel .

The weight $w(ji)$ was calculated below:

$$w(i,j) = exp\left(-\frac{|B(i) - B(i)|^2}{h^2}\right) \tag{2}$$

Whereas $h$ denotes the filtering parameter, which controls the function of exponential decay and $B(i)$ as:

$$B(i) = \frac{1}{|R(i)|} \sum_{q \in R(i)} v(q) \tag{3}$$

Here $R(i) \subseteq \Omega$ and refers to the squared area of pixel $i$; $|R(i)||R(p)|$ represents pixel counts in the region $R$.

**3.1.2. Contrast Enhancement**

The histogram modification method is applied in the AGCWD method [20]. This method integrates the gamma correction and HE techniques. Gamma correction is a transform-based histogram modification model. This model uses variable parameters (gamma). As a pre-determined value has been applied for all images, the Gamma Correction model gives inconsistent outcomes for all images. HE has problems with under and over-development. By combining both methods and using a weighting function, the AGCWD method removed the disadvantages of either HE or gamma correction. This model that utilizes a normalized cumulative density function, utilizes Gamma Correction. The transform-based gamma correction in its most simple form is computed by

$$T(L) = L_{max}(L/L_{max})^{\gamma} \tag{4}$$

Now $L_{max}$ denotes the highest mammography image intensity of the input $L$ represents pixel intensity within the input image and is converted as $T(L)$ according to Eq. (4). The Gamma curves are given through a value higher than one having the precise opposite influence as this produced with a value lower than 1. If the contrast is unswervingly influenced by Gamma Correction, dissimilar photos will present similar intensity variations owing to the constant value. It is assumed to compute the $pdf$ as shown.

$$pdf(L) = \frac{n_L}{M^*N} \tag{5}$$

Now $n_L$ refers to the amount of pixels that have pixel value $L$ and $M^*N$ implies total pixel counts in the image. The $cdf$ based $pdf$ is expressed as

$$cdf(L) = \sum_{k=0}^{L} pdf(k) \tag{6}$$

The CDF-based transformation is provided as demonstrated

$$T(L) = cdf(L)^* L_m \tag{7}$$

If $the\ pdf$ contains the maximum entropy and is the most uniform, the model's transformation techniques are the identity line. To reimburse for the limitations of typical development methods, a model that strikes cooperation amongst excellent visual quality and inexpensive computing costs should be advanced. Some numerical information might be lost in the corrected sub-histograms decreasing the influence of the development. Therefore, the reduced the gamma value the superior the size of the correction.

$$T(L) = L_{max}(L/L_{max})^{(1-cdf(L))} \tag{8}$$

The AGC method permits for a slow growth in smaller intensities despite the fact that disregards a larger reduction in higher intensities. The Weighted Distribution (WD) has been additionally applied to slightly fine-tune the histogram and stop negative removes from developing. The next is how the WD work is written:

$$pdf_w(L) = pdf_{max}\left(\frac{pdf(L) - pdf_{min}}{pdf_{max} - pdf_{min}}\right)^{\alpha} \tag{9}$$

Now $\alpha$ denotes the tuning parameter, $pdf_{max}$ refers to the maximal $pdf$ of the histogram, and $pdf_{min}$ denotes lowest $pdf$ according to Eq. (9), the altered $CDF$ is estimated as

$$cdf_w(L) = \sum_{=0}^{L_{max}} pdf_w(L) / \sum pdf_w L \quad (10)$$

The sum of $pdf_w$ is computed as shown

$$\sum pdf_w = \sum_{l=0}^{L_{max}} pdf_w(L) \quad (11)$$

The gamma parameter according to $cdf$ of Eq. (8) is modified as demonstrated

$$\gamma = 1 - cdf_W(L) \quad (12)$$

The AGCWD flowchart model for the dark image is applied as input. The function of weighted distribution can smooth the fluctuating events and lower the gamma corrections above amplification.

## *3.2. Multi-Model Feature Extraction*

Besides, the presented EMSFF-KOSDDMO model designs a multi-scale feature fusion of EfficientNetB0 with MBConv4 and EfficientNetV2-B0 with MBConv6 for the extraction of feature process to capture rich, multi-scale features with efficient computational performance.

### 3.2.1. EfficientNetB0 with MBConv4

EfficientNet comprises the CNN family and is intended for efficient and superior image classification [21]. EfficientNetB0 is the basic method of the EfficientNet family EffecientNetB0 has proved important achievement in tasks. Its architecture comprises a scaling model for constant size resolution scaling and scaling for advanced performance. This method is adjusted for image classification tasks. The EfficientNetB0 structure can be derived from MBConv constituents in addition to squeeze-and-excitation (SE) blocks. It combines depth separable convolutional to decrease cost of computational by $k^2$, whereas $k$ denotes size of kernel. The structure additionally contains inverted residual blocks to reduce the trainable parameter counts. Resolution, depth, and width are scaled for suitable pattern scaling by the EffecientNetB0 method. During compound scaling, a compound co-efficient $\phi$ has been applied as stated in Eq. (13).

$$d = \alpha\psi, w = \beta\psi, r = \gamma\psi \quad (13)$$

Whereas $\geq 1$, $\beta \geq 1$, and $\gamma \geq 1$ are scaling components; $\phi$ relies on model scaling.

The Swiss function can be described by Eq. (14):

$$fswiss(x) = \frac{\chi}{1 + e^{-\beta(x)}} \quad (14)$$

The last layer of the EfficientNetB0 structure named the head has the responsibility for forecasting the last classification output according to the features extracted from the previous layers.

EfficientNet-B0, an advanced DL method, uses the MBConv block, especially MBConv4, as a main building block for its effective framework. MBConv4 is a depth-wise separable convolution block by an expansion factor of 4 that meaningfully decreases computational cost while preserving higher precision. This block integrates point-wise convolutions, depth-wise convolutions, and SE mechanisms to enhance feature extraction and improve representative capacity. By utilizing MBConv4, EfficientNetB0 attains an outstanding exchange amongst performance and size of the model, making it appropriate for tasks needing effective calculation and higher precision, like object detection and image classification in resource-constrained settings. Fig. 2 signifies the framework of EfficientNetB0 with MBConv4.


**Fig. 2.** Framework of EfficientNetB0 with MBConv4

#### 3.2.3. EfficientNetV2-B0 with MBConv6

This study selects EfficientNet-V2 for training [22]. This model is an enhanced version of EffcientNet-V1, inspired by determining that carefully balanced network resolution, depth, and width can result in improved performance. The major block of EfficientNet-V1 is the squeeze-and-excitation (SE) optimization and Inverted Residual Block (MBConv) block, while EfficientNet-V2 utilizes Fused-MBConv that substitutes the depth-wise Conv$3x3$and development Conv$1x1$ in MBConv by an individual regular Conv$3x3$ to substitute some initial phases of the MBConv. The $input$ parameter is in the layout of HWC and it contains a convolution (Conv) $1x1$ (kernel dimension) block for growing feature size, a Depthwise Conv $3x3$ $(or\ 5x5)$ block performing convolutions while all kernel simply performs convolution on the network. In addition to this, there are furthermore 3 other changes amongst EfficientNet-V2 and EfficientNet-V1 that are emphasized as:

1. EfficientNet-V2 supports a small expansion ratio for MBConv since small expansion ratios normally lead to low accessing memory overhead.
2. EfficientNet-V2 prioritizes utilizing small $3x3$ kernel dimensions, though it contains advanced layers to compensate for the small receptive field generated by the small kernel size.
3. EfficientNet-V2 removes the final stride-l stage applied in the new EfficientNet, probably due to its important parameter dimensions and the complexity related to memory access.

By significantly assessing the above-mentioned points, EffcientNet-V2 is considered suitable for being an outstanding CNN approach. Consequently, most of the investigators have utilized EffcientNet-V2 in their application. EfficientNetV2-B0, a sophisticated version of the EfficientNet family, combines the MBConv6 block, which provides improved efficacy and performance in comparison with previous modifications. The MBConv6 block, with an expansion ratio of 6, utilizes an integration of point-wise convolutions and depth-wise separable convolutions, enhancing the ability of the model to learn composite characteristics with decreased computational complexities. The principal innovation in MBConv6 is its usage of a more aggressive expansion of channels throughout the convolution procedure, permitting it to take more complex models while maintaining model parameters under control. These outcomes improved faster training and accuracy of the model, making EfficientNetV2-B0 with MBConv6 mainly suitable for higher-performance tasks.

### *3.3. Classification using TabNet method*

For the detection process, the TabNet system has been developed. TabNet is a new DL approach designed for numerical tables, incorporating the interpretability of decision trees with the flexibility and power of neural networks (NNs) [23]. It outshines for its tactical usage of a learnable attention mechanism (AM) to dynamically choose features, making it expert at processing higher-dimensional data and composite connections amongst features. The TabNet framework is tailored for processing data over a series of decision stages, permitting sequential attention to various feature subsets at all stages. This architecture permits the method to concentrate on the most related features to make predictions, similar to how a human predictor may approach a dataset by focusing on various features consecutively.

The TabNet presented contains many important modules namely Feature Transformer, Attentive Transformer, Output layer, and Decision step. The Attentive Transformer has the responsibility to select the features being applied at all decision stages. It uses masking to the input features, establishing which of these are related according to the present state of the method and the data as shown in Eq. (15).

$$M = Sparemax\left(FC\big(BN(D_{last})\big)\right) \qquad (15)$$

Whereas $M$ refers to the mask produced for feature selection, FC signifies a fully connected layer, BN characterizes batch normalization used for the decision output from the preceding stage. $D_{last}$ and Sparsemax represent a sparsity-inducing function, which guarantees just a features subset is designated.

When the Attentive Transformer chooses features, the Feature Transformer makes a more complex representation of these features. It is important to capture composite interactions and relationships amongst the chosen features as presented in Eq. (16).

$$H = ReLU\left(FC\big(BN(XoM)\big)\right) \tag{16}$$

Now, $H$ characterizes the converted features, $X$ refers to the input feature set, $\circ$ $and$ signifies element-to-element multiplication using the mask $M$, guaranteeing just designated features are handled. The architecture of TabNet contains numerous decision stages, all containing a Feature Transformer and an Attentive Transformer. These stages iteratively enhance the concentration of the model and data representation, providing piecewise to the last prediction that can be represented in Eq. (17).

$$D_i = ReLU\left(FC\big(BN(H_{i-1})\big)\right) \tag{17}$$

Whereas, $D_i$ denotes decision output at step $i$, by $H_{i-1}$ is the output of the Feature Transformer from the preceding stage. This recursive procedure permits the sequential improvement of information. Afterward, pass through the selected decision step counts, and the outputs are aggregated to create the last prediction of the model. This aggregation seizures the complete study presented by the method through every step as exposed in Eq. (18).

$$Y = \sum_{i} \quad FC(D_i) \tag{18}$$

Whereas, $Y$ denotes final prediction, aggregating the contributions ($D$) from all decision stages over FC layers.

### *3.4. Hyperparameter Tuning using DMOA*

At last, the DMOA algorithm alters the hyperparameter values of TabNet system and outcomes in greater performance of classification. Basically, the DMOA is a bioinspired optimizer method, which outlines stimulation from the social and behavior communications of DMs and smaller mongooses in Africa [24]. One of the most important aspects of their behavior is cooperative foraging, while DM foraging in flocks, sharing and communicating information about sources of food. This cooperative hunting approach guarantees that the group effectively discovers their setting to discover and use sources of food. This part surrounds members to guarantee that all facilitate the safety inside groups, improving the model's robustness and flexibility. Furthermore, the separation of labor between DMs, like grooming, foraging, and sentinel duty, is flexible and dynamic, permitting the group to answer efficiently to alter the state of the environment. By combining these cooperative and social behaviors, the DMOA is tailored for exploiting and exploring the searching region successfully balancing exploitation with exploration. The model was easily communicated.

In general, swarm intelligence-based optimizer methods initiate with random initialization of the family or herd. Formerly, based on the diversification and intensification guidelines, all solutions, or individuals, begin collecting avoiding global optimal targets, or the alpha female. In the same way, the optimization of DMOA starts with the initialization of the mongoose's candidate population ($M$), as in Eq. (19). Population was built randomly in the provided issues $LL$ (Lower Limit) and $UL$ (Upper Limit).

$$M = \left[m_{1.1}\ m_{1,2}\ \ldots\ m_{1,y-1}\ m_{1,y}\ m_{2.1}\ m_{2,2}\ \ldots\ m_{2,y-1}\ m_{2,y}\ \vdots \vdots\ m_{p.z}\ \vdots \right.$$
$$\left. \vdots\ m_{x.1}\ m_{x,2}\ \ldots\ m_{x,y-1}\ m_{x,y}\ \right] \tag{19}$$

Whereas $M$ characterizes the meeting of the current population's mongoose candidates, which are randomly generated utilizing Eq. (19), specifies the location of $z^{th}$ size of $p^{th}$ population, $x$ specifies the total dimension of population, and $y$ denotes problem dimension.

The predefined function of MATLAB $unifrnd$ in Eq. (20), has been applied as a device for making a sequence of evenly distributed randomly produced numbers. Besides, $D$ mentions the decision variable counts of the mongooses, whereas $UB$ and $LB$ represent problems with upper and lower boundaries, individually.

$$m_{p,z} = unifrnd(LB, UB, D) \tag{20}$$

The optimal outcome in all training is deliberated as the best-fit outcome thus far consequently, afterward the initialization ends, the fitness degree, defined in Eq. (21). The members using the maximum fit possibility are selected as an alpha female.

$$\alpha = \frac{Fit_n}{\sum_{n=1}^{x} \quad Fit_n} \tag{21}$$

The real number of mongooses in $\alpha$ accepted to $x - bb$, while the babysitter counts $bb$. The vocalization of leading female $P$ represents the vivacious sound. The solution's upgraded model was calculated with Eq. (22):

$$S_{n+1} = S_n + \emptyset * P \tag{22}$$

Now, $\emptyset$ represents randomly distributed numbers. The sleep mound $\partial$ was calculated for all repetitions Eq. (23).

$$\partial_n = \frac{Fit_{n+1} - Fit_n}{max\{|Fit_{n+1}, Fit_n|\}} \tag{23}$$

The average of $\partial_n$ was computed in Eq. (24):

$$\gamma = \frac{\sum_{n=1}^{x} \quad \partial_n}{x} \tag{24}$$

| **Algorithm 1:** Pseudo-code of DMOA |
|---|
| Initialization: |
| Start |
| Initialization DMOA parameters: population $x$ and babysitter counts $bb$. |
| Set $x = x - bb$. |
| Set exchange of babysitter $P_{bb}$. |
| for $i = 1: Max_i$ |
| Compute all mongoose's individual Fit. |
| Set time Counter T. |
| Compute $\alpha = \frac{Fit_n}{\Sigma_{n=1}^{\chi} Fit_n}$ |
| Compute the candidate's food location $S_{n+1} = S_n + \emptyset * P$ |
| Estimate novel Fit of $S_{n+1}$. |
| Estimate sleeping mound $\partial_n = \frac{Fit_{n+1} - Fit_n}{max\{|Fit_{n+1}, Fit_n|\}}$ |
| Calculate the average of $\partial_n . \gamma = \frac{\Sigma_{n=1}^{\chi} \partial_n}{\chi}$ |
| Calculate movement vector $\overrightarrow{M_v} = \sum_{n=1}^{x} \quad \frac{S_n \times \partial_n}{S_n}$ |
| Exchange of babysitter if $T \geq P_{bb}$ and set. |
| Initialization $bb$ location using and compute fitness $Pit_n \leq \alpha$. |
| Simulate; next location $S_{n+1} = \{S_n - CVM * \emptyset * rand[S_n - \overrightarrow{M_v}]\ is\ \theta_{n+1} < \theta_n Exploration\ S_n + CVM * \emptyset * rand[S_n - \overrightarrow{M_v}]\ \ else\ Exploration$ |
| Updated best solution. |
| end for |
| Return the best solution. |
| End |

The DMOA system originates a fitness function (FF) for achieving the heightened solution of classifier. It expresses a positive number for representing the better results of candidate solution. The reduction of the classifier rate ratio is evaluated as FF. Its formulation is set in Eq. (25).

$$fitness(x_i) = ClassifierErrorRate(x_i)$$

$$= \frac{no.of\ misclassified\ instances}{Total\ no.of\ instances} * 100 \qquad (25)$$

## 4. Result Analysis and Discussion

The investigational validation of EMSFF-KOSDDMO approach is examined under the Knee Osteoarthritis database with severity grading [25]. This database holds 8950 images under five grades as depicted in Table 1. Fig. 3 depicts the sample images.

**Table 1** Details of the dataset

| **KL Grading** | **Descriptions** | **Images** |
|---|---|---|
| "Grade 0" | "Healthy knee image" | 3500 |
| "Grade 1 (Doubtful)" | "Doubtful joint narrowing with possible osteophytic lipping" | 1500 |
| "Grade 2 (Minimal)" | "Definite presence of osteophytes and possible joint space narrowing" | 2500 |
| "Grade 3 (Moderate)" | "Multiple osteophytes, definite joint space narrowing, with mild sclerosis" | 1200 |
| "Grade 4 (Severe)" | "Large Osteophytes, significant joint narrowing, and severe sclerosis" | 250 |
| **Total Number of Images** | | **8950** |

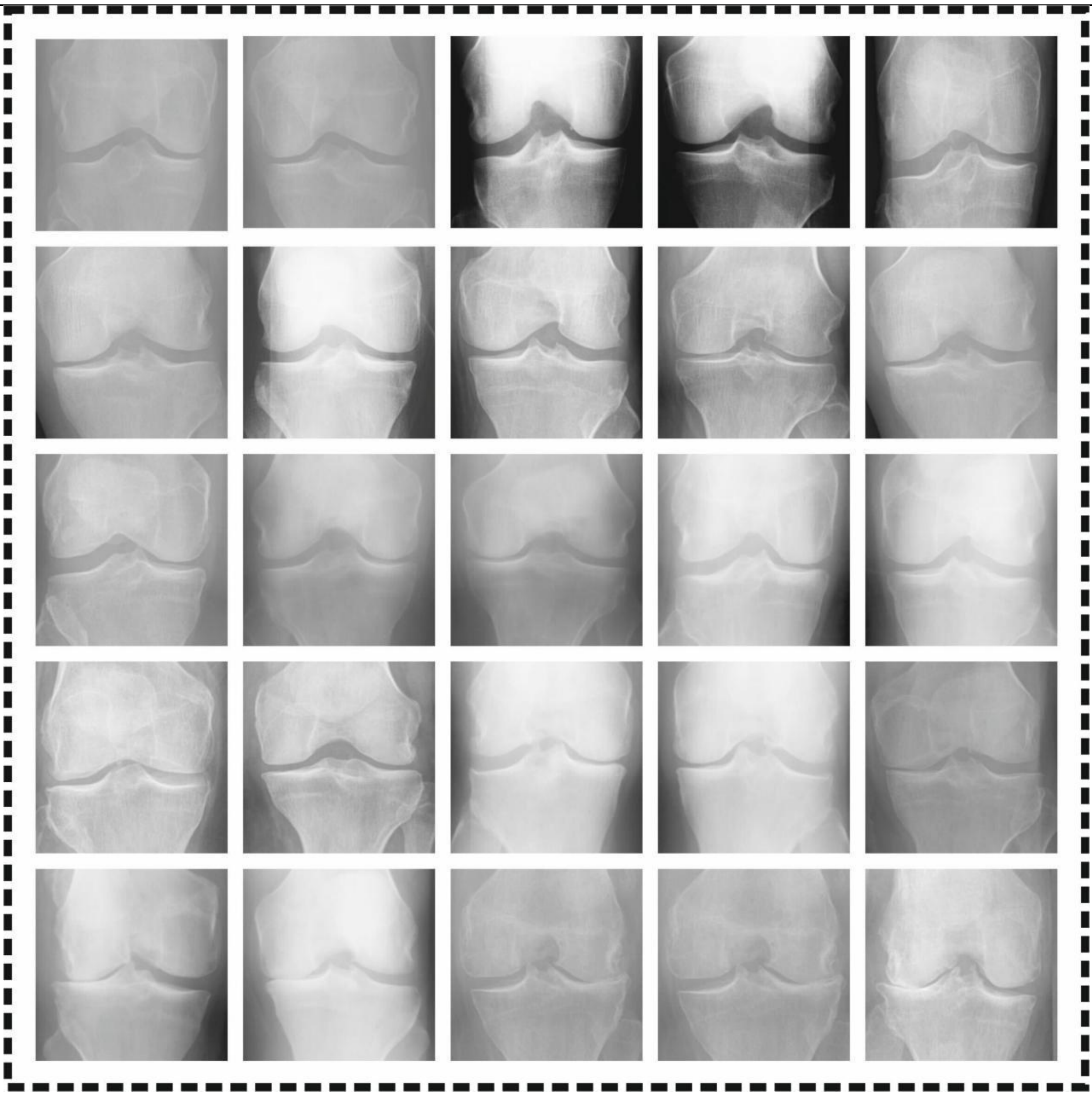

**Fig. 3.** Sample Images

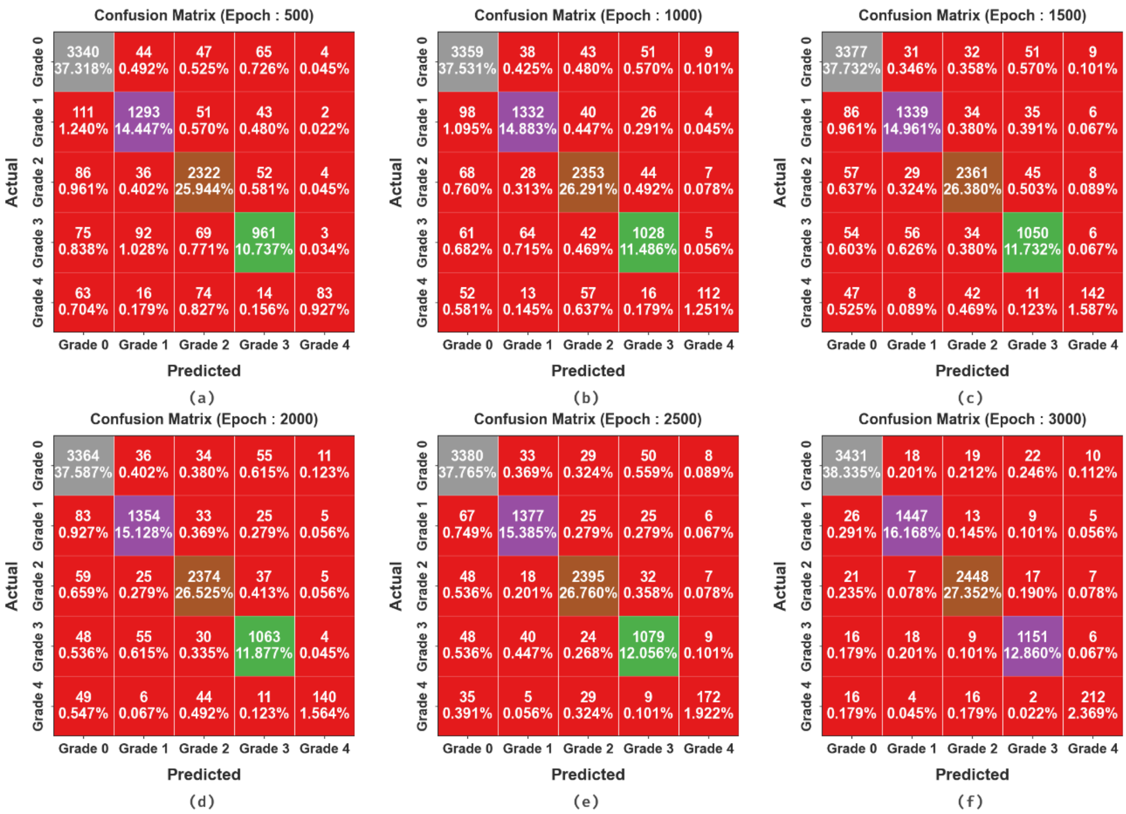


**Fig. 4.** Confusion matrices of EMSFF-KOSDDMO technique (a-f) Epochs 500-3000

Fig. 4 shows the confusion matrices formed by EMSFF-KOSDDMO model below several epochs. The outcomes identify that the EMSFF-KOSDDMO system has proficient detection of every classes exactly.

The classifier results of EMSFF-KOSDDMO algorithm are determined below different epochs in Table 2 and Fig. 5. The outcomes shows that the EMSFF-KOSDDMO system correctly recognized all the instances. On 500 epochs, the EMSFF-KOSDDMO technique offers an average $accu_y$ of 95.75%, $prec_n$ of 87.98%, $reca_l$ of 77.56%, $F1_{score}$ of 80.37%, and MCC of 78.74%. Also, on 1000 epochs, the EMSFF-KOSDDMO algorithm attains an average $accu_y$ of 96.58 $prec_n$ of 89.09%, $reca_l$ of 81.87%, $F1_{score}$ of 84.39%, and MCC of 82.67%. In addition, on 1500 epochs, the EMSFF-KOSDDMO method provides an average $accu_y$ of 96.96 $prec_n$ of 90.05%, $reca_l$ of 84.90%, $F1_{score}$ of 86.97%, and MCC of 85.21%. Furthermore, on 2500 epochs, the EMSFF-KOSDDMO system obtains an average $accu_y$ of 97.56 $prec_n$ of 91.82%, $reca_l$ of 88.58%, $F1_{score}$ of 90.02%, and MCC of 88.48%. At last, on 3000 epochs, the EMSFF-KOSDDMO approach delivers an average $accu_y$ of 98.83 $prec_n$ of 95.30%, $reca_l$ of 94.63%, $F1_{score}$ of 94.96%, and MCC of 94.18%.

**Table 2** Classifier results of EMSFF-KOSDDMO system under different epochs

| **Class labels** | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F1_{score}$ | $MCC$ |
|---|---|---|---|---|---|
| **Epoch - 500** | | | | | |
| Grade 0 | 94.47 | 90.88 | 95.43 | 93.10 | 88.56 |
| Grade 1 | 95.59 | 87.31 | 86.20 | 86.75 | 84.10 |
| Grade 2 | 95.32 | 90.60 | 92.88 | 91.72 | 88.47 |
| Grade 3 | 95.39 | 84.67 | 80.08 | 82.31 | 79.70 |
| Grade 4 | 97.99 | 86.46 | 33.20 | 47.98 | 52.87 |
| **Average** | **95.75** | **87.98** | **77.56** | **80.37** | **78.74** |
| **Epoch - 1000** | | | | | |
| Grade 0 | 95.31 | 92.33 | 95.97 | 94.12 | 90.26 |

| | | | | | |
|---|---|---|---|---|---|
| Grade 1 | 96.53 | 90.31 | 88.80 | 89.55 | 87.47 |
| Grade 2 | 96.32 | 92.82 | 94.12 | 93.47 | 90.91 |
| Grade 3 | 96.55 | 88.24 | 85.67 | 86.93 | 84.96 |
| Grade 4 | 98.18 | 81.75 | 44.80 | 57.88 | 59.74 |
| **Average** | **96.58** | **89.09** | **81.87** | **84.39** | **82.67** |
| **Epoch - 1500** | | | | | |
| Grade 0 | 95.90 | 93.26 | 96.49 | 94.85 | 91.48 |
| Grade 1 | 96.82 | 91.52 | 89.27 | 90.38 | 88.48 |
| Grade 2 | 96.86 | 94.33 | 94.44 | 94.38 | 92.20 |
| Grade 3 | 96.74 | 88.09 | 87.50 | 87.79 | 85.91 |
| Grade 4 | 98.47 | 83.04 | 56.80 | 67.46 | 67.97 |
| **Average** | **96.96** | **90.05** | **84.90** | **86.97** | **85.21** |
| **Epoch - 2000** | | | | | |
| Grade 0 | 95.81 | 93.37 | 96.11 | 94.72 | 91.27 |
| Grade 1 | 97.01 | 91.73 | 90.27 | 90.99 | 89.20 |
| Grade 2 | 97.02 | 94.39 | 94.96 | 94.68 | 92.60 |
| Grade 3 | 97.04 | 89.25 | 88.58 | 88.92 | 87.21 |
| Grade 4 | 98.49 | 84.85 | 56.00 | 67.47 | 68.24 |
| **Average** | **97.07** | **90.72** | **85.18** | **87.36** | **85.71** |
| **Epoch - 2500** | | | | | |
| Grade 0 | 96.45 | 94.47 | 96.57 | 95.51 | 92.58 |
| Grade 1 | 97.55 | 93.48 | 91.80 | 92.63 | 91.17 |
| Grade 2 | 97.63 | 95.72 | 95.80 | 95.76 | 94.12 |
| Grade 3 | 97.35 | 90.29 | 89.92 | 90.10 | 88.58 |
| Grade 4 | 98.79 | 85.15 | 68.80 | 76.11 | 75.95 |
| **Average** | **97.56** | **91.82** | **88.58** | **90.02** | **88.48** |
| **Epoch - 3000** | | | | | |
| Grade 0 | 98.35 | 97.75 | 98.03 | 97.89 | 96.53 |
| Grade 1 | 98.88 | 96.85 | 96.47 | 96.66 | 95.99 |
| Grade 2 | 98.78 | 97.72 | 97.92 | 97.82 | 96.98 |
| Grade 3 | 98.89 | 95.84 | 95.92 | 95.88 | 95.24 |
| Grade 4 | 99.26 | 88.33 | 84.80 | 86.53 | 86.17 |
| **Average** | **98.83** | **95.30** | **94.63** | **94.96** | **94.18** |

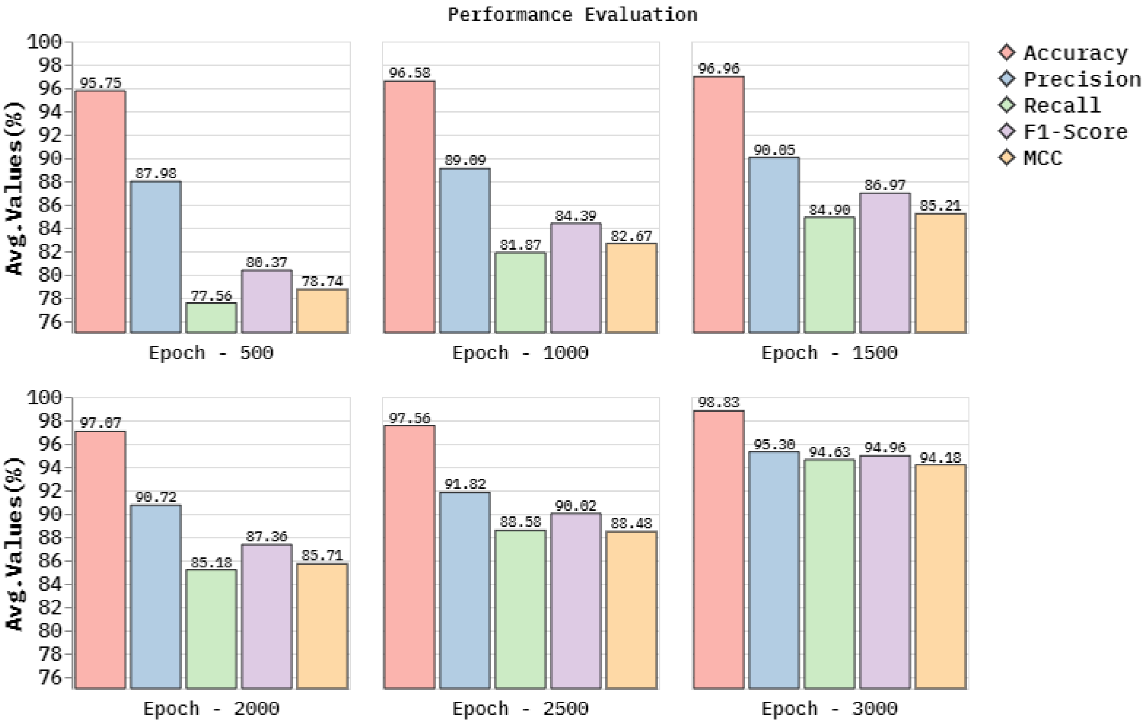


**Fig. 5.** Average of EMSFF-KOSDDMO technique under distinct epochs

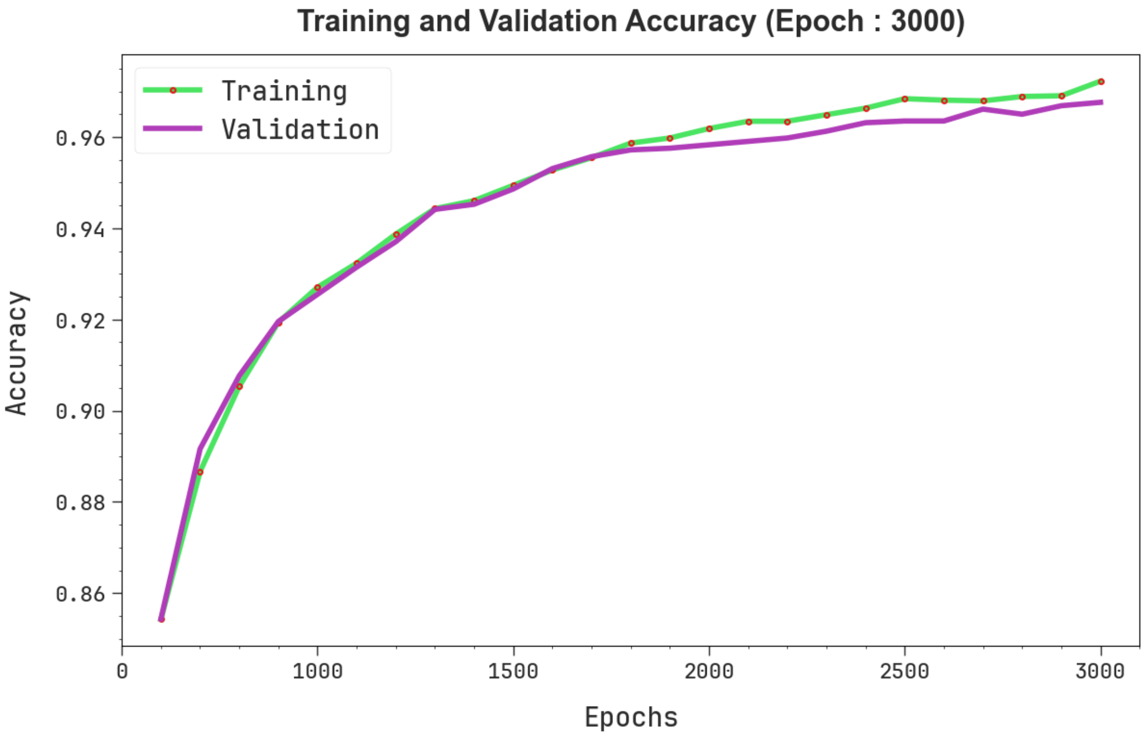


**Fig. 6.** $Accu_y$ curve of EMSFF-KOSDDMO technique on Epoch 3000

In Fig. 6, the training (TRAN) $accu_y$ and validation (VALD) $accu_y$ analysis of EMSFF-KOSDDMO methodology on epoch 3000 is illustrated. The $accu_y$ study is computed across the interval of 0 to 3000 epochs. The figure emphasizes that both analysis exhibited an increasing trend which informed the capacity of EMSFF-KOSDDMO approach with superior performance across manifold iterations. Additionally, the both $accu_y$ remains nearer across the epochs, which specifies lesser overfitting and exhibits maximum outcomes of the EMSFF-KOSDDMO algorithm.

In Fig. 7, the TRAN loss and VALD loss curves of the EMSFF-KOSDDMO system under epoch 3000 are shown. The values of loss are computed across an interval of 0 to 3000 epochs. It is represented that the both values exemplify a decreasing trend to balance a trade-off among data fitting and simplification. The constant decrease promises higher outcomes of the EMSFF-KOSDDMO approach.

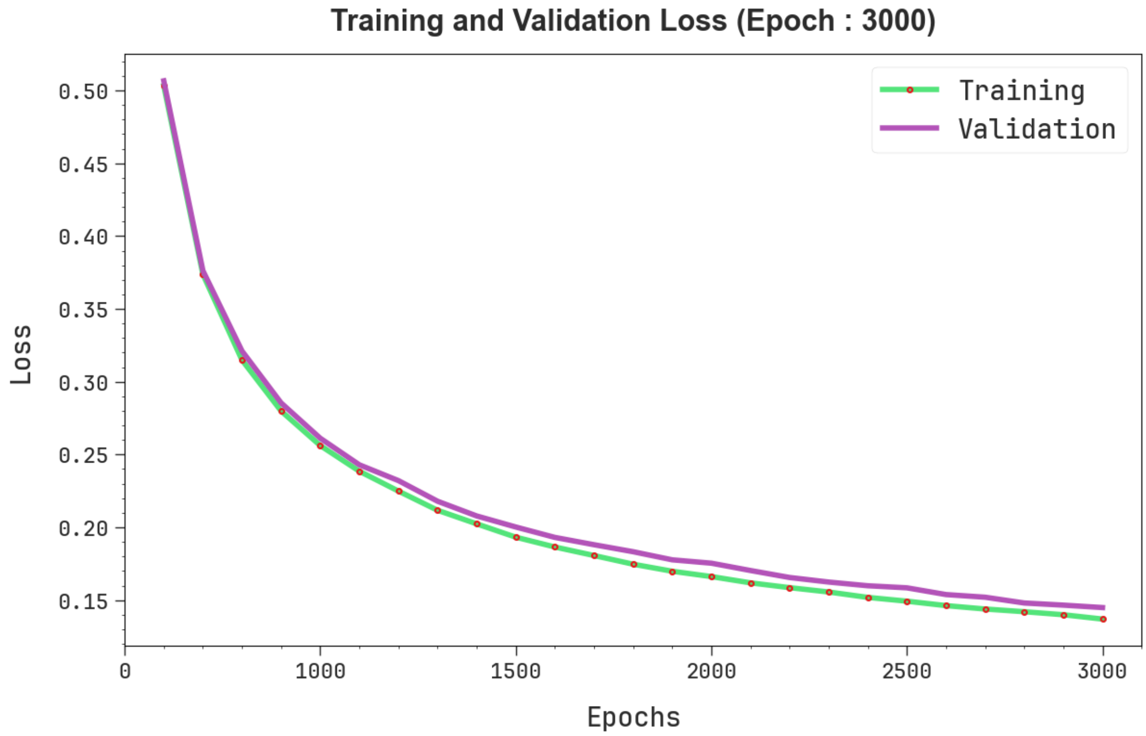


**Fig. 7.** Loss graph of EMSFF-KOSDDMO technique below Epoch 3000

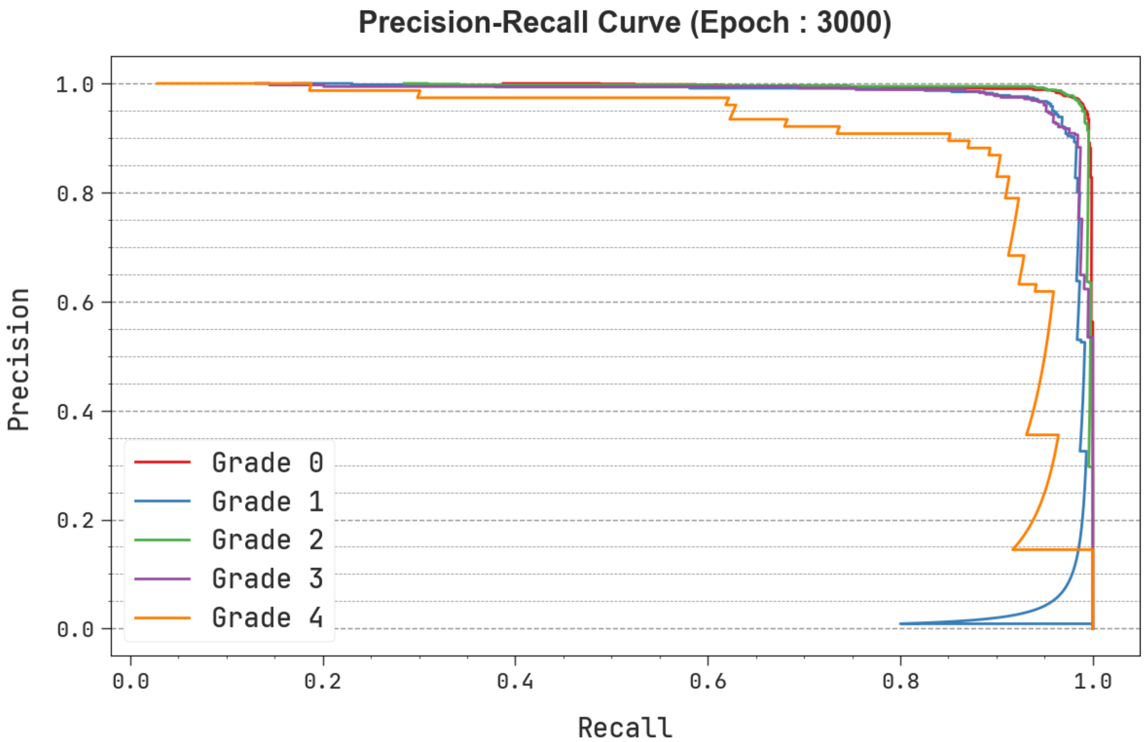


**Fig. 8.** PR analysis of EMSFF-KOSDDMO methodology under Epoch 3000

In Fig. 8, the precision-recall (PR) graph results of the EMSFF-KOSDDMO technique under epoch 3000 offer clarification into its results by plotting Precision beside Recall for every class label. The figure demonstrates that the EMSFF-KOSDDMO technique always achieves maximal PR analysis over dissimilar classes. The stable rise among each class label represents the effectiveness of the EMSFF-KOSDDMO technique.

In Fig. 9, the ROC analysis of the EMSFF-KOSDDMO algorithm on epoch 3000 is considered. The outcomes recommend that the EMSFF-KOSDDMO methodology completes superior ROC results over every class. This dependable tendency of highest ROC analysis across manifold class labels of the EMSFF-KOSDDMO technique on forecasting class labels.

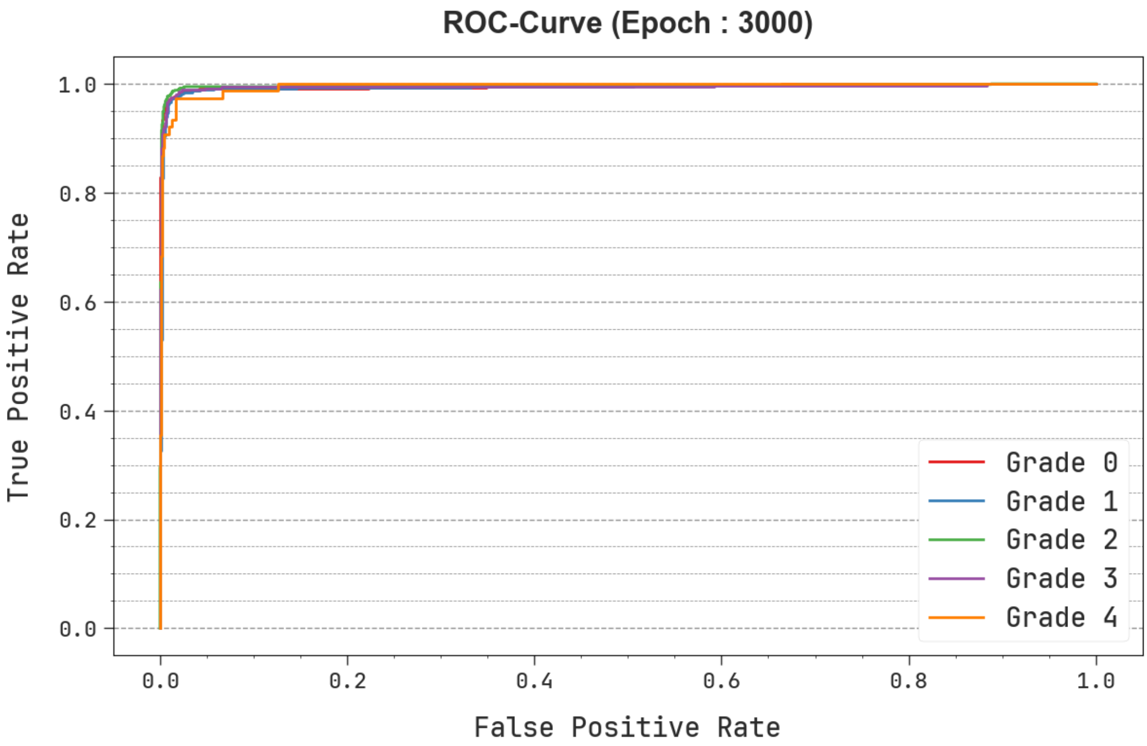


**Fig. 9.** ROC curve of EMSFF-KOSDDMO technique on Epoch 3000

Table 3 and Fig. 10 present the comparative examination of the EMSFF-KOSDDMO algorithm [26-28]. The outcomes suggest that the proposed EMSFF-KOSDDMO approach has achieved a greater performance $accu_y$ of 98.83%, $prec_n$ of 95.30%, $reca_l$ of 94.63%, and $F1_{score}$ of 94.96%. Meanwhile, the existing systems VGG-16, ResNet-101, EfficientNetb7, DenseNet121, 2DConvNet, CNN, and SVM algorithms have gained the worst results.

**Table 3** Comparative outcomes of EMSFF-KOSDDMO system with existing algorithms

| **Model** | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F1_{score}$ |
|---|---|---|---|---|
| EMSFF-KOSDDMO | 98.83 | 95.30 | 94.63 | 94.96 |
| VGG-16 Method | 96.55 | 91.16 | 88.19 | 86.97 |
| ResNet-101 Model | 94.58 | 92.36 | 87.15 | 86.40 |
| EfficientNetb7 | 95.91 | 88.69 | 90.00 | 86.22 |
| DenseNet121 | 83.39 | 90.99 | 85.67 | 91.30 |
| 2DConvNet | 79.39 | 87.91 | 90.59 | 85.12 |
| CNN Classifier | 79.58 | 90.45 | 87.69 | 93.50 |
| SVM Method | 83.56 | 87.88 | 86.97 | 92.04 |

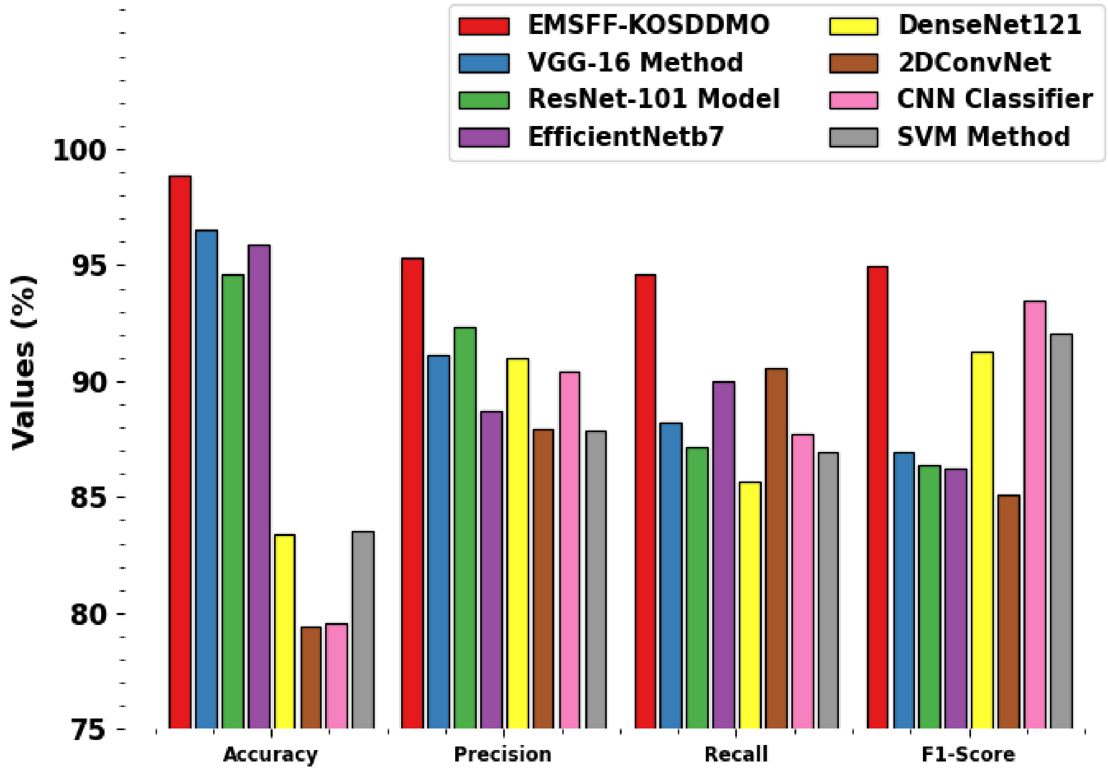


**Fig. 10.** Comparative outcomes of the EMSFF-KOSDDMO method with existing models

In Table 4 and Fig. 11, the time outcome of EMSFF-KOSDDMO system with existing methodologies are illustrated. The outcomes suggest that the EMSFF-KOSDDMO technique gets maximum outcomes. Dependent on time, the EMSFF-KOSDDMO system delivers lower time of 5.74min while the VGG-16, ResNet-101, EfficientNetb7, DenseNet12, 2DConvNet, CNN, and SVM approaches reach better time of 26.62min, 16.41min, 9.55min, 15.72min, 19.49min, 22.87min, and 29.60min, correspondingly.

**Table 4** Time result of EMSFF-KOSDDMO methodology with recent methods

| **Technique** | **Time (min)** |
|---|---|
| EMSFF-KOSDDMO | 5.74 |
| VGG-16 Method | 26.62 |
| ResNet-101 Model | 16.41 |
| EfficientNetb7 | 9.55 |
| DenseNet121 | 15.72 |
| 2DConvNet | 19.49 |
| CNN Classifier | 22.87 |
| SVM Method | 29.60 |

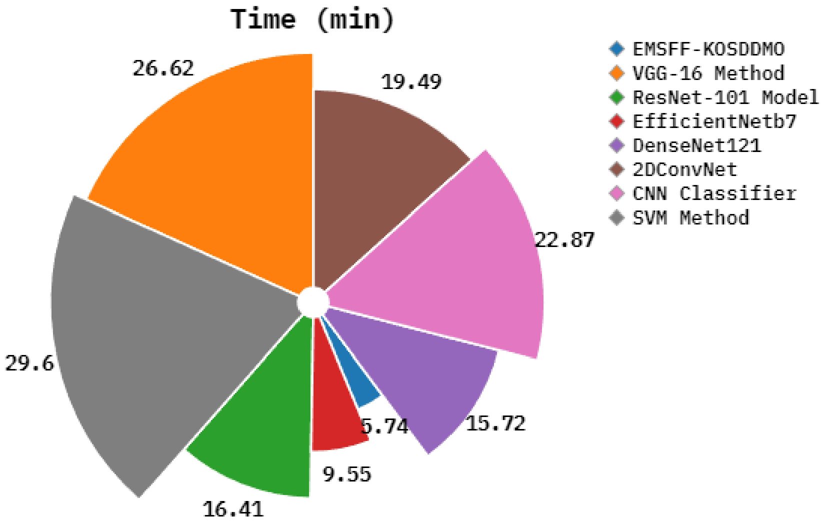


**Fig. 11.** Time outcome of EMSFF-KOSDDMO technique with existing models

## 5. Conclusion

In this manuscript, we offer an EMSFF-KOSDDMO system. The main intention of EMSFF-KOSDDMO method is to enhance the recognition of knee osteoarthritis severity using medical image data. Initially, the EMSFF-KOSDDMO technique applies NLM filtering for noise removal and AGCWD for contrast enhancement to enhance the knee X-ray image quality. Besides, the presented EMSFF-KOSDDMO model designs a multi-scale feature fusion of EfficientNetB0 with MBConv4 and EfficientNetV2-B0 with MBConv6 for the extraction of feature process to capture rich, multi-scale features with efficient computational performance. For the detection process, the TabNet system has been exploited. At last, the DMOA algorithm alters the hyperparameter value of TabNet system and outcomes in greater performance of classification. A wide sort of experimentations was conducted to authorize the performance of EMSFF-KOSDDMO system under various evaluation measures. The simulation results suggested that the EMSFF-KOSDDMO system emphasized improvement over other existing techniques.

**Data Availability Statement**

The data that support the findings of this study are openly available in the Kaggle repository at https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity, reference number [25].

**Funding Statement**

This research received no external funding.

**Conflict of Interest**

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

**Ethics approval**

This article does not contain any studies with human participants performed by any of the authors.

**Consent to Participate**

Not applicable.

**Informed Consent**

Not applicable.

**References**

1. Antony, J., McGuinness, K., Moran, K. and O'Connor, N.E., 2017. Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks. In Machine Learning and Data Mining in Pattern Recognition: 13th International Conference, MLDM 2017, New York, NY, USA, July 15-20, 2017, Proceedings 13 (pp. 376-390). Springer International Publishing.

2. Chen, P., Gao, L., Shi, X., Allen, K. and Yang, L., 2019. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. Computerized Medical Imaging and Graphics, 75, pp.84-92.
3. Pongsakonpruttikul, N., Angthong, C., Kittichai, V., Chuwongin, S., Puengpipattrakul, P., Thongpat, P., Boonsang, S. and Tongloy, T., 2022. Artificial intelligence assistance in radiographic detection and classification of knee osteoarthritis and its severity: a cross-sectional diagnostic study. Eur Rev Med Pharmacol Sci, 26(5), pp.1549-1558.
4. Yang, J., Ji, Q., Ni, M., Zhang, G. and Wang, Y., 2022. Automatic assessment of knee osteoarthritis severity in portable devices based on deep learning. Journal of Orthopaedic Surgery and Research, 17(1), p.540.
5. Cueva, J.H., Castillo, D., Espinós-Morató, H., Durán, D., Díaz, P. and Lakshminarayanan, V., 2022. Detection and classification of knee osteoarthritis. Diagnostics, 12(10), p.2362.
6. Katz, J.N., Arant, K.R. and Loeser, R.F., 2021. Diagnosis and treatment of hip and knee osteoarthritis: a review. Jama, 325(6), pp.568-578.
7. Abdullah, S.S. and Rajasekaran, M.P., 2022. Automatic detection and classification of knee osteoarthritis using deep learning approach. La radiologia medica, 127(4), pp.398-406.
8. Thomas, K.A., Kidziński, Ł., Halilaj, E., Fleming, S.L., Venkataraman, G.R., Oei, E.H., Gold, G.E. and Delp, S.L., 2020. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. Radiology: Artificial Intelligence, 2(2), p.e190065.
9. Ribas, L.C., Riad, R., Jennane, R. and Bruno, O.M., 2022. A complex network based approach for knee Osteoarthritis detection: Data from the Osteoarthritis initiative. Biomedical Signal Processing and Control, 71, p.103133.
10. Abd Alla, I.H., Mohammed, F.M., Al-din, S. and Abdulbaqi, A.S., 2025. Integrating a Secure and Low-Cost WSN Layer with Medical Cloud Computing for Medical Image Transmission. Fusion: Practice and Applications, (1), pp.35-5.
11. Mehdi, T., Nasser, Y., Aloui, S. and Jennane, R., 2025. Deep learning based approach combining shape and texture features for knee osteoarthritis prediction from X-ray images. Biomedical Signal Processing and Control, 101, p.107172.
12. Rani, S., Memoria, M., Almogren, A., Bharany, S., Joshi, K., Altameem, A., Rehman, A.U. and Hamam, H., 2024. Deep learning to combat knee osteoarthritis and severity assessment by using CNN-based classification. BMC Musculoskeletal Disorders, 25(1), p.817.
13. Kadu, R. and Pawar, S., 2025. Advanced Bi-CNN for Detection of Knee Osteoarthritis using Joint Space Narrowing Analysis. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 7(1), pp.80-90.
14. Jain, R.K., Sharma, P.K., Gaj, S., Sur, A. and Ghosh, P., 2024. Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. Multimedia Tools and Applications, 83(3), pp.6925-6942.
15. Fang, Y., Liu, L., Yang, Q., Hao, S. and Luo, Z., 2025. A new method for early diagnosis and treatment of meniscus injury of knee joint in student physical fitness tests based on deep learning method.
16. Sivakumar, C., Deepti, K., Revanth, R.S., Gowthami, M. and Geethika, C.S., 2024, March. Implementation of Knee Osteoarthritis Detection and Severity Prediction Using Convolution Neural Network. In 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1998-2005). IEEE.
17. Segura, F.M., Segura, F.P., Zudaire, M.P.L. and Segura, F.V., 2025. Advances in Artificial Intelligence for automated knee osteoarthritis classification using the IKDC system. European Journal of Orthopaedic Surgery & Traumatology, 35(1), pp.1-7.
18. Karegoudra, S., Veeresha, R.K. and Yendapalli, V., 2024, November. Deep Learning Approaches for Detecting and Predicting Knee Osteoarthritis Severity. In 2024 2nd International Conference on Recent Advances in Information Technology for Sustainable Development (ICRAIS) (pp. 108-111). IEEE.
19. Taassori, M. and Vizvári, B., 2024. Enhancing Medical Image Denoising: A Hybrid Approach Incorporating Adaptive Kalman Filter and Non-Local Means with Latin Square Optimization. Electronics, 13(13), p.2640.
20. Srinivas, A., Prasad, V.V. and Kumari, B.L., Segmentation of Mammogram Images Using Optimized Kernel Fuzzy AGCWD Based Level Set Method.
21. Iqbal, A., Jaffar, M.A. and Jahangir, R., 2024. Enhancing Brain Tumour Multi-Classification Using Efficient-Net B0-Based Intelligent Diagnosis for Internet of Medical Things (IoMT) Applications. Information, 15(8), p.489.
22. Zhao, Z., Bakar, E.B.A., Razak, N.B.A. and Akhtar, M.N., 2024. Corrosion image classification method based on EfficientNetV2. Heliyon, 10(17).
23. Khattak, A., Zhang, J., Chan, P.W., Chen, F., Almaliki, A.H. and Mongina Matara, C., 2024. Estimating Turbulence Due to Low-Level Wind Shear in Airport Runway Zones Using TabNet-SHAP Framework. IEEE Access, 12, pp.101325-101341.
24. Hesham, O.M., Attia, M.A. and Mekhamer, S.F., 2024. Enhancement of AVR system performance by using hybrid harmony search and dwarf mongoose optimization algorithms. Scientific Reports, 14(1), p.27177.
25. https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity
26. Ahmed, R. and Imran, A.S., 2024. Knee Osteoarthritis Analysis Using Deep Learning and XAI on X-rays. IEEE Access.
27. Mohammed, A.S., Hasanaath, A.A., Latif, G. and Bashar, A., 2023. Knee osteoarthritis detection and severity classification using residual neural networks on preprocessed x-ray images. Diagnostics, 13(8), p.1380.
28. Liu, L., Chang, J., Zhang, P., Ma, Q., Zhang, H., Sun, T. and Qiao, H., 2023. A joint multi-modal learning method for early-stage knee osteoarthritis disease classification. Heliyon, 9(4).