FRACTURE FUSION:
REVOLUTIONIZING THE RECOGNITION OF BONE FRACTURES WITH METAMAG EFFICIENCY APPROACH

S. Rajeashwari* , Dr. K. Arunesh

Department of Computer Science, Sri S. Ramasamy Naidu Memorial College,
Affiliated to Madurai Kamaraj University, Sattur, India
*Corresponding author: S. Rajeashwari (rajeeragavan83@gmail.com)

Abstract Bone fractures are common in diabetic patients and can result in several musculoskeletal conditions. Both type 1 and type 2 diabetes substantially increase the risk and severity of bone fractures. Prompt treatment and management of diabetes and its complications are crucial to mitigate this serious complication. Detection and diagnosis in its early stage can reduce the challenging conditions in treatment. Traditional image processing techniques like digital-geometric analysis, entropy measures, and gray-level co-occurrence matrices have been used for automated bone fracture detection. However, these detection methods rely neither on healthy controls nor diabetic-affected patients. Only few studies focused on detecting fractures in diabetic patients. The rising prevalence of diabetic ankle fractures made the study emphasize the development of a fracture detection model based on the Meta Magnify (MetaMag) efficiency model. The proposed model involves the Lower Extremity Radiographs (LERA) dataset, which consists of image samples of normal and abnormal lower extremities of the body, such as the hip, ankle, knee, and foot. Pre-processing involves a one-hot encoding method that handles the missing data and represents categorical variables as numerical values. Further, the classification is performed using the MetaMag efficiency model, incorporated with MetaMag scaling and unified normalization. Further, the efficiency of the proposed model is analyzed by comparing it with conventional EfficientNet and another model. Finally, the proposed work's performance is analyzed using evaluation measures such as accuracy, precision, recall and F1-score. The results indicate the improved efficiency of the model.

Keywords: fracture, Lower Extremity Radiographs dataset, diabetes, Deep Learning, radiograph images, EfficientNet.

1. Introduction

Among other parts of the body, the knee is considered the most complex joint that involves many daily activities. A high prevalence of knee injuries occurs due to twisting movements and sudden changes of direction [7]. This creates chances of knee damage and other risk factors leading to severe impact on the patient's lifestyle. Approximately one in eight patients has diabetes and undergoes treatment for rotational ankle fractures. With this, complications of ankle fracture fixation in patients with Diabetes Mellitus (DM), after surgery vary between 26% and 47% [20]. Several researchers have also identified that an ankle injury may trigger the process of Charcot neuroarthropathy. These higher complication rates can cause bone deformity, loss, and joint destruction. The most affected areas damaged due to an injury are the patellofemoral, ligaments,

cartilage, and meniscus. In addition, the data analysis has resulted that a large cohort of 58 748 patients who undergo ankle fracture fixation in New York discovered that 12.5% were diabetic, and 14.6% of patients resulted in complicated diabetes [10]. Moreover, the widely used methods involved in detecting lesions in the knee part are Magnetic Resonance Imaging (MRI) and X-ray copies [3]. The results produced by these methods are promising, but there is still a need to develop new equipment and research. So, in recent times, AI has emerged as the significant opinion of specialists that assist in providing non-invasive tools, and low-complexity and low-cost instruments [11]. These methods enable the system to extract the patterns from the input data and map the relationships among the input variables and outcomes. Thus, these new technologies tend to efficiently identify knee abnormalities and diagnosing methods at their early stages to avoid higher consequences of disease in patients. Although these techniques effectively detect and interpret fractures in DM patients, they lack high detection accuracy due to the irregularity and lucidity in the input sample images.

On the contrary, several studies investigated the prediction of knee fractures in diabetes patients by using Machine Learning (ML) and Deep Learning (DL) algorithms [1, 31]. Hence, the considered study [25] implements Convolutional Neural Network (CNN), to perform the detection of abnormality on lower extremity radiographs. The lower extremity includes the range of abnormalities in hip, knee, ankle, and foot radiographs. This study's larger dataset comprises almost 93 455 input samples of lower extremity radiographs of several body parts. These samples are labeled as normal and abnormal at the initial interpretation by the attending radiologist. The CNN is pre-trained with 161-layer densely connected to achieve improved accuracy in the process of classification. The performance of the study was analyzed by using three different models such as pre-trained ResNet-101, DenseNet-161 [30], and ResNet-50. Further, an extensive random hyperparameters search for each model is performed. The motive of the study is to provide increased accuracy in the classification tasks. This is done by augmenting the dataset by using MURA radiographs, this tends to optimize the efficacy of the model. From analysis, it is found that the DenseNet-161 produced better diagnostic accuracy. In the other aspects, the intimated study [2] applies the detection of Anterior Cruciate Ligament (ACL) using the DL model. The model involves the customized 14-layer ResNet-14 structure of CNN and six directions. This is done by involving real-time data augmentation and hybrid class balancing. Three classes are classified: ruptured tears, partial and healthy. Initially, the data pre-processing undergoes three steps and after the steps, the three classes are raised. The original version-I residual ResNet-18 in the classification model is modified into ResNet-14 network architecture. Here, the Batch Normalization (BN) is added after the CNN model and previous to the activation function Rectified Linear Unit (ReLu) [29]. The fine-tuned hyperparameters are being used that provide a huge impact on the effectiveness of the method. The outcomes of the study projected better outcomes in terms of accuracy, specificity, sensitivity, F1-score, precision, and

AUC. However, the techniques failed to produce improved classification accuracy and enhanced input sample images to perfectly interpret the affected region [19].

The present study aims to further optimize the automatic detection of fractures by using a DL model with radiographic images. The MetaMag efficiency model using MetaMag Scaling along with Unify Normalization is proposed in the present study, which tends to significantly increase the detection of fractures and classifies whether the input is fractured or non-fractured. Input from the LERA dataset is first passed into the pre-processing stage, where the one-hot encoding method is applied. This method endeavors to handle missing values and generates efficient features for classification. Then the pre-processed data are fed into the train-test phase, where the train data are used for pre-training the classifier. In the classification process, the MetaMag efficiency model undergoes MetaMag scaling that uniformly scales all the dimensions of resolution, width and depth for procuring improved performance. It systematically analyses the model scaling and identifies the balancing network using a simple yet highly effective compound coefficient. This work focuses on improving the practical efficiency of the traditional EfficientNet model by using the unified normalization that reduces the computational loss and inexpensively fine-tuning at higher resolution. It eventually increases the size of the image and aids in obtaining finer details of the input image. This helps the classifier distinguish the input images into two categories: normal as 0 and abnormal as 1. Thus, the efficiency of the proposed model is evaluated by using performance measures.

1.1. The main contributions of the study

- To efficiently classify the normal and abnormalities in the input LERA dataset, to detect fractures in the lower extremities of the human body.
- To implement MetaMag scaling and Unify Normalization approaches to precisely analyze the attributes and improve classification accuracy.
- To evaluate the model's efficacy by involving performance measures: accuracy, recall, precision, and F1-score.
- To compare the proposed MetaMag efficiency model with other conventional algorithms to project the effectiveness of the proposed system.
- To develop MetaMag efficiency model for improved bone fracture classification accuracy and efficiency, as well as plans to create automated systems to assist clinicians in diagnosis and treatment planning.

1.2. Organization of the paper

The remaining parts of this paper are organized as follows. Section 2 deliberates the review of conventional works with the problems identified by analysis of several studies. Section 3 expounds on the projected procedures with the proposed flow, algorithms, and their mathematical derivations. Subsequently, section 4 presents the results attained by

the proposed and conventional models. The overall study is concluded in Section 5 with future suggestions.

1.3. Motivation of the research

Patients with both type 1 and type 2 diabetes have a significantly increased risk of bone fractures compared to those without diabetes. Diabetes can impair bone quality and fracture healing, leading to a higher risk of complications like delayed union, non-union, or prosthetic joint formation. Identifying and managing bone fragility in diabetic patients is an emerging challenge that requires more attention, as current osteoporosis and diabetes guidelines do not adequately address this issue. Improving the understanding and management of bone health in diabetes is crucial to mitigate this serious complication. Hence, the proposed model utilises LERA for effective classification process.

2. Literature review

The analysis of various studies on fracture detection using different strategies and the methodologies and problem identification for specific studies are also deliberated.

The human knee joints are the main and complex joints present in the human body that maintain weight and offer flexible movements of the body. It bears the excess load and is thus highly prone to injuries. So, detecting knee injuries as early as possible is important to avoid complications and provide appropriate treatments.

The study [14] involved the prediction of Knee Osteoarthritis (KOA) using the ML-based approach. The study has applied a multidisciplinary Osteoarthritis Initiative (OAI) database collected through self-reported data on joint symptoms, physical activity indexes, disability and function, physical examination data, and questionnaire data. Initially, the data pre-processing has been done by implying data imputation to tackle the missing values. Then, the feature selection was done by integrating the output of six feature selection algorithms, three embedded techniques, one wrapper, and two filter algorithms. Whereas, the ML-based techniques like Logistic Regression (LR), k-Nearest Neighbor (KNN), Random Forest (RF), Naive Bayes (NB), Decision Tree (DT), XGBoost and SVM have been evaluated for validating their sustainability been utilized to solve the classification issues. The better accuracies produced by these models have been identified and found that the SVM model has performed better, producing an accuracy of 74.07%. Even though the model has been reliable, the predictive capacity has to be improved predominantly.

Knee abnormalities are mostly due to hard injury or osteoarthritis that greatly impact the patient's health. Generally, the MRI plays a vital part in detecting the biochemical and morphologic features that provide an in-depth understanding of patterns. So, the suggested study [23] has MRI-based studies to conduct the identification of lesion severity

in the ACL, meniscus, bone marrow, and cartilage. A three-dimensional CNN has been developed to identify the Region of Interest (ROI) and then grade the abnormalities. At first, the segmentation was performed by using two V-Net architectures under two consecutive steps. From analysis, it has been consolidated that the study has produced improved specificity, sensitivity, and multiclass lesion severity staging in several tissues of the knee. In addition, the generalizability of the model has to be improved, and the assessment of lateral and medial ligaments has to be considered. On the other hand, the intimated study [13] has relied on the detection of abnormalities and classification automatically using Musculo-Skeletal Disorders Network (MSDNet). These methods have been an ensemble of CNN that integrates the features of several CNN models to improve the performance of abnormality classification. A boundary detection algorithm has been developed to predict the ROI to facilitate enhanced detection of anomalies. The MSDNet is the combination of both AlexNet and ResNet18 structures. Firstly, the global features have been produced from the AlexNet by directly feeding the original input data, whereas the local features have been generated by the ResNet18 model. The overall accuracy produced by the MSDNet model is 82.69%. Among aged people, the main factor for fracture [6, 24, 29] is due to the reduction of bone density. A low-cost diagnostic technology is important in identifying osteoporosis in its initial stage. So, the suggested study [15] has analyzed osteoporosis using X-ray radiography to predict the essential components and categorize it into osteoporosis, osteopenia, and normal. The study has implemented three CNN architectures namely, ResNet18, Xception, and Inceptionv3 models. This ensemble method has implied a fuzzy rank-based fusion of classifiers by considering the two different factors. A fuzzy ranking-based approach has been applied, which has been exposed to two distinct non-linear processes. After implementation, the study's outcomes have shown that the study has produced a classification accuracy of 93.5%. The accuracy has been hindered due to overlapping cells or insufficient picture quality that made complexity in classifying the images effectively.

The advancements in radiological technologies have improved the treatment of various diseases. But, when compared with a huge number of fractural patients, the number of radiologists is insufficient. This makes radiologists astounded by the large amount of medical image data. Hence, the imitated study [12] has deployed a backbone network by applying dilated convolutions to detect the fractured thigh region. The DL method known as Dilated Convolutional Feature Pyramid Network (DCFPN) has been used, in which stage 1 has been adopted to extract the features from the original image. It has been insisted that the dilated convolutional kernel could gain more information from the extended receptive field. The FPN structure has been comprised of five feature maps. The Region Proposal Network (RPN) has been developed to generate the region proposal that shares the convolutional feature maps. Thus, the output is an image with a predicted bounding box. The radiologists have used the technique of Computer-Aided Diagnosis (CAD) to diagnose the fractures [8, 22] on bones, which minimizes their

difficulties. Thus, the suggested study [16] has involved classification using a Crack Sensitive Convolutional Neural Network (Crack-Net) to identify the sensitive fracture lines on human bones. This paper clearly explains the two different stages of discovering the fracture [4, 21, 26]. Initially, Faster R-CNN, which is Faster Region with CNN, was deployed. This method has been performed to identify 20 types of bone regions and fractures [27] using Crack-Net in the collected X-ray copies. The results have shown that from the total of 1052 copies 526 copies are fractured copies.

Further, the study has produced an accuracy rate of 90.11% and an F-measure of 90.14% of the x-ray copies. In radiographs, the method of Guided Anchoring (GA) Faster R-CNN has been used to identify and locate the fractures in hand [28]. This GA method has resulted in improved, accurate, and effective anchor generation. It has eventually increased the network's performance and saved computing energy. In this system, the Feature Pyramid Network (FPN) method has been used to detect small fractures [5, 9] such as knuckles and fingertips joints and others. Additionally, the implementation of balanced loss (L1) has been applied to adapt imbalanced learning tasks. The result of this system has shown that among 3067 HF dataset X-ray copies, 2453 are training data and 614 are testing data. The accuracy of the dataset has been achieved to be 97%-99% with an Average Precision (AP) of 70.7%. This System has accomplished all the other conventional methods for identifying HF.

Problem identification

- The study has involved the detection of fractures using X-ray images. Though the system has produced a better detection rate, the classification accuracy can be prominently improved by applying different algorithms [16].

- The risk factors accompanied by knee osteoarthritis have been involved in the study using DL models. The accuracy produced by the study has been identified to be 74.07%. It can be further improved to support radiologists in finding the complexities [14].

- Binary classification of lower extremity fracture has been performed in the study and produced limitation of producing generalizability in detecting the abnormalities. Efficient methods can be applied to detect the fracture [25].

3. Proposed methodology

DM is a metabolic disorder that increases the chance of interfering with bone formation and fracture risk. This leads to the impairment of fracture healing and several other common features that affect the bone. DL techniques greatly impact the medical domain and lead to advancements in the detection of abnormalities that help in affording early diagnosis of diseases. There is still a lack of studies investigating the association between DM and fracture risk in patients. The possible solutions to fracture risk should
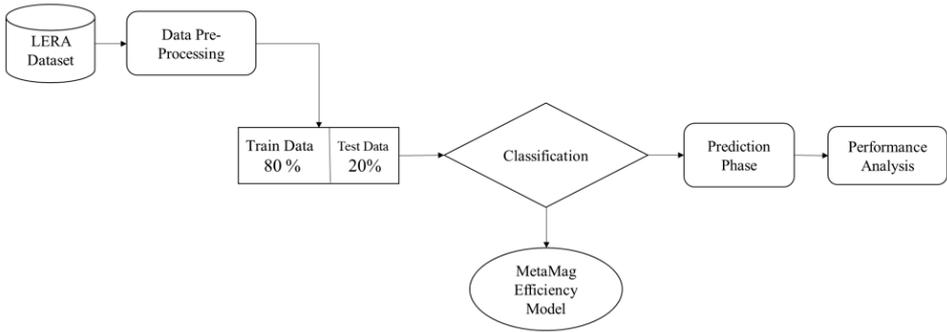
Fig. 1. Illustrative diagram of the overall methodology

be addressed at an early stage to avoid the severity of risk in patients with diabetes. Spontaneous calcaneal fractures without obvious trauma may occur in diabetic patients sometimes accompanied by DFU. With this intention, the initial phase concentrated on detecting the foot ulcer in DM patients. The study implemented a Deep Convolutional Neural Network (DCNN) based on the Xception model to classify healthy and DFU skin images. The DCNN-based Xception classifier was integrated with Residual Linearly Clamped Layers (RLCL) comprising minimum detached convolution layers. Further, the input images are optimized by using image enhancement techniques such as Histogram equalization, Adaptive filter, and Gamma correlation. Then, the efficiency of the proposed system is evaluated based on the performance measures, namely precision, F1-score, recall, and accuracy, to validate the performance of the proposed model with existing algorithms. Though the study has proclaimed improved efficiency. It is noteworthy that patients manifesting systematic signs of diabetic foot infection cause fractures or dislocations of the ankle or foot. With this regard, it is also significant to address the challenges faced by the diabetic patients with lower limb amputations. So, the present work focussed on detecting and classifying the normal and fractured bone classes by using the MetaMag efficiency model. This method tends to reduce the problems related to high-risk factors and efficiently contributes towards risk reduction and management. The overall process involved in the proposed technique is demonstrated in Figure 1.

The input from the LERA dataset (see Section 3.1) is first passed into the pre-processing stage, where the one-hot encoding process is applied. This method tends to handle the missing values and generates efficient features for classification. Then the pre-processed data are fed into the train-test phase, where the train data are used for pre-training the classifier. Further, the classification is performed by a MetaMag efficiency classifier that involves MetaMag scaling and a unified normalization process that supports enhancing the performance of the proposed method. The classifier classifies

Tab. 1. Class Distribution of LERA Dataset.

Samples	Hip	Foot	Ankle	Knee
Abnormal images	3	36	36	99
Normal images	91	12	285	435
Total images	94	348	321	534

the input images into two categories: normal as 0 and abnormal as 1. Wherein the prediction phase validates the classifier's efficiency by using test data and analyses by using performance measures.

Association of diabetes with fractures

- DM type 1 and type 2 affect several people worldwide and are characterized by hyperglycemia. The traditional impediments of DM are microvascular complications like neuropathy, nephropathy, and retinopathy. Whereas the macrovascular complications include CVD (Cardiovascular Disease). The researchers have also found that diabetes affects the bones of DM patients with increased chances of fracture due to impaired bone quality. Further, the fracture risk in diabetes patients can be described by possible cofounders, diabetes type, and fracture site.
- Type 1 DM is related to a modest reduction of bone mineral density. Type 2 DM increases the chance of affecting bone health in its advanced phases of disease. The biomechanical characteristics of bone and bone architecture are negatively impacted by chronic inflammation, Advanced Glycation End products (AGE), hyperglycemia, and insulinopenia.
- Several methods are used to evaluate bone quality in DM, including the diagnosis based on X-ray images, MRI images, Grayscale images, Red Green Blue (RGB) images, and radiography images.

3.1. Dataset description

The dataset used in the proposed method is LERA [17], which covers the broad range of joints and bone abnormalities of lower extremity areas of the human body. The dataset is considered a diverse-natured dataset due to its collection over a wide range of time, from 2003 to 2014. This LERA dataset comprises anomalous and standard image dissemination and sample images of hip, ankle, knee, and foot bones. This dataset has been accumulated by HIPAA complaint that compiled data from almost 182 patients who have undergone radiographic examination at Standard University Medical Centre. A total of 1297 normal and abnormal images of lower extremities have been presented in the dataset. Table 1 shows the class distribution of the LERA dataset.

The LERA dataset is one of the benchmark musculoskeletal radiograph image datasets and has been applied in the proposed approach for producing a relatively improved

a

b
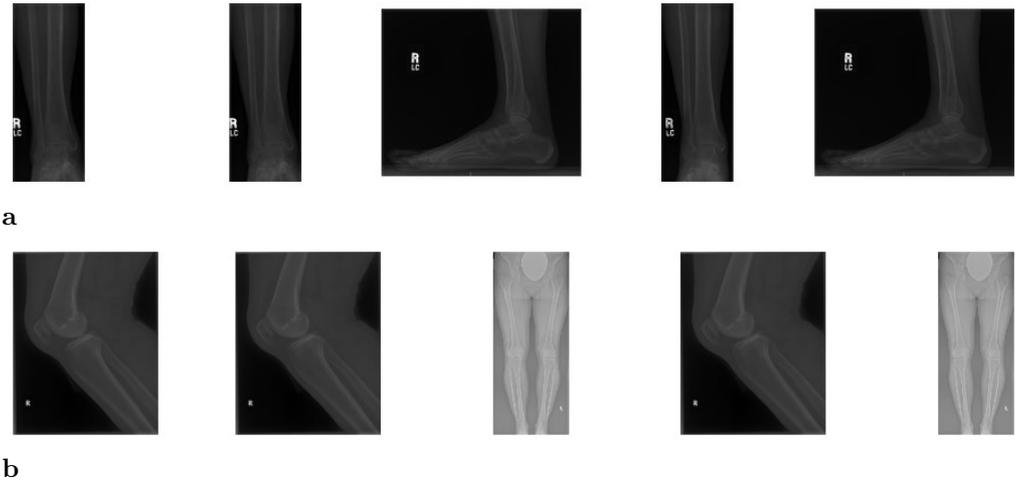
Fig. 2. LERA Dataset – (**a**) normal and (**b**) abnormal image samples.

degree of classification accuracy. Moreover, the interpretation in binary classification is distinguished in a way that abnormal as "1" and normal as "0". The abnormal categorization denotes that the radiograph consists of either fractures or any other abnormalities. Meanwhile, in normal categorization, the radiographs represent that the image is normal. The Figure 2 presents the sample images of the LERA dataset.

3.2. Pre-processing techniques

The image pre-processing method is applied in the input image to predominantly enhance the radiographic image's eminence, the edges that denote the possible fractures. This study's proposed method involves a one-hot encoding-based pre-processing approach to rectify the missing data issues.

One-Hot Encoding The one-hot encoding is a type of encoding method and is considered to be the most popular target encoding technique. The main advantage of this strategy is that it is a sparse vector, which is used in calculating the similarities or distances between the features for efficient classification. Here, one element is set to 1, and all other elements are set to 0. Contradictory to the other existing algorithms, the one-hot encoding method treats all missing values as a new class. This tends to mitigate the interference with data structure in the simulation.

3.3. Train and test split of data

The input LERA dataset consisting of normal and abnormal images of the lower extremities of the human body is split into train and test datasets. The splitting of data is done with 80% of train data and 20% of test data. The splitting of input data is such that training data gains more than two-thirds of the entire data. The training dataset is used in training the classifier employed in classifying the normal and abnormal images. The test data are applied to compute the performance measures.

3.4. Classification

3.4.1. EfficientNet model

The conventional EfficientNet is a kind of NN which uses the compound scaling method to produce better system performance. These existing models target to improve the performance and computational efficiency by subsiding the Floating Point Operations Per Second (FLOPS) and several parameters. Scaling up mechanisms involved in EfficientNet are Neural Architecture Search (NAS) and compound scaling. Initially, the baseline network is designed by performing NAS, a method used to automate the design of neural networks. It efficiently optimizes both efficiency and accuracy as measured on a FLOPS basis. The two parts present in EfficientNet are created using a baseline with NAS and compound scaling to increase the performance. Compared with other state-of-arts models, the EfficientNet significantly reduces the computational resources required to train the classifier. The scaling method involved in EfficientNet has shown uniform scaling across multiple dimensions. This could be more efficient when applied to a highly versatile architecture to improve the effectiveness of the model. When combined with CNN, the EfficientNet involves a scaling approach and achieves significant output in the performance.

3.4.2. MetaMag efficiency classifier

The MetaMag efficiency classifier is deployed in the proposed method, where the network architecture involves a new scaling model known as MetaMag scaling. The other existing CNNs randomly scale the network dimensions like resolution, dimension, and width. The MetaMag efficiency model uniformly scales the entire image with a fixed scaling coefficient. This tends to enhance the efficiency and accuracy of classification. In the classification process, the MetaMag efficiency model undergoes MetaMag scaling that uniformly scales all the dimensions of resolution, width and depth for procuring improved performance. It systematically analyses the model scaling and identifies the balancing network using a simple yet highly effective compound coefficient.

MetaMagnify scaling

The scaling factor, denoted as ϕ, allows for adjustments in the depth of the network. When ϕ is increased, the model becomes deeper and more robust, enhancing its capability to extract complex features. This is advantageous for tasks that demand sophisticated feature extraction, such as intricate pattern recognition in images or nuanced language understanding. Conversely, reducing ϕ results in a shallower model. This can be advantageous for simpler tasks or scenarios where computational resources are restricted. Shallow models are effective for straightforward classification tasks or when rapid inference speed is crucial. Furthermore, smaller values of ϕ facilitate faster training and reduce memory requirements. This makes them particularly suitable for environments where efficiency in model development and deployment is prioritized.

Unify normalization

The use of Unify Normalization offers a way to maintain the benefits of Batch Normalization (BN) while addressing its challenges with large activation memory requirements due to the need for sizable batch sizes. This is particularly relevant in memory-intensive AI accelerators that rely on local memory for enhanced speed and energy efficiency, despite tighter memory constraints. Additionally, our approach aims to preserve BN's normalization advantages while circumventing its regularization effects when they prove counterproductive. To adapt the EfficientNet architecture effectively, it is essential to adjust the initial scaling operations within the network. This ensures that scaling factors play a significant role in shaping the overall network structure. Furthermore, modifying batch normalization layers to accommodate variations in network width and depth is crucial for maintaining effective normalization during training.

Besides, this work focuses on improving the practical efficiency of the traditional EfficientNet model by using the unified normalization that reduces the computational loss and inexpensively fine-tuning at higher resolution. It eventually increases the size of the image and aids in obtaining finer details of the input image. This helps the classifier distinguish the input images into two categories: normal as 0 and abnormal as 1. The input data from the training dataset is fed into the input layer of the MetaMag efficiency classifier and then to the MetaMag scaling layer. By using this layer, the finer details of radiographic images are obtained that precisely classify the abnormalities found in the bone. The process involved in the MetaMag efficiency model is shown in Figure 3.

The MetaMag efficiency model uses the MetaMag scaling method that involves a series of fixed factors to scale the dimension of the network in a uniform manner based on resolution, depth, and width. The building block i is defined as a function of $A_{i+1} = B_i(A_i)$, where B_i denotes the operator and A_i represents the input tensor, and A_{i+1} is the output tensor. Thus, the CNN, denoted symbolically as n, is characterized by different layers as given in equation (1),

$$n = B_m^{q_m} \odot \cdots \odot B_2^{q_2} \odot B_1^{q_1}(A_1) = \odot_{i=1,\ldots,m} B_i(A_1), \tag{1}$$
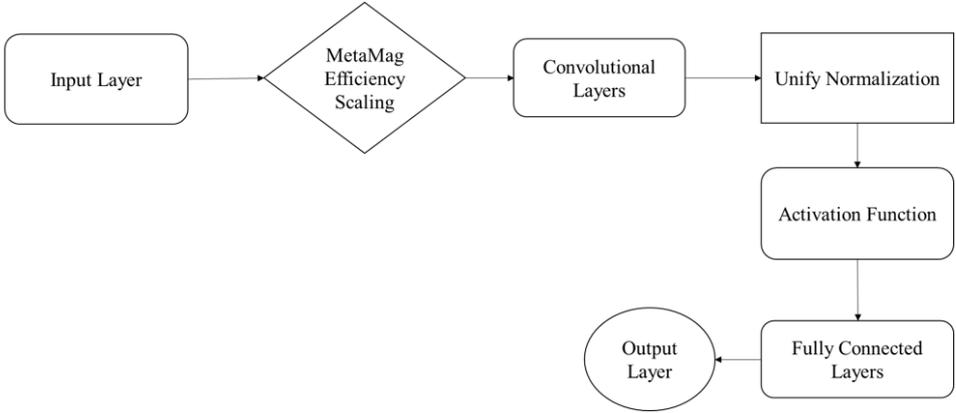
Fig. 3. Flow diagram of the proposed MetaMag efficiency classifier.

where $\odot$ is the Hadamard product, that is, the element-wise multiplication of two matrices, and the superscript q_i denotes the hyperparameter vector of B_i. This epitomizes the architecture of building block i, which is not able to be determined from training. Further, m signifies the number of layers present in the network. Further on, the hyperparameter matrix q with a building block defined in CNN is shown in equation (2),

$$n = \odot_{i=1,\dots,m} B_i^{q_i} (A_{C_i, D_i, H_i, W_i}). \tag{2}$$

The proposed modified EfficientNet model aims to resolve the optimization problem formulated in equation (3),

$$q_{\text{optimum}} = \arg \max_q \text{Accuracy} \left(n(q)(A_{C_i, D_i, H_i, W_i}) \right), \tag{3}$$

where q is the matrix of hyperparameters of the whole network, formed by vectors q_i of the subsequent operators B_i. The denotation $n(q)$ underlines the dependency of the network on its parameters. Therefore, the result of a search procedure of the modified EfficientNet model is the optimal hyperparameter matrix q. The architecture of the proposed modified EfficientNet model is displayed in Figure 4. In this structure, the convolution pooling layers consist of extracted features from the input radiographic images and `conv` blocks that process the feature maps. Further, the unified normalization is performed at the end of the network.

The modified `conv` layer i is defined by the function $Y_i = B_i(X_i)$, in which B_i is the operator, X_i denotes the input tensor, and Y_i represents the output tensor. The tensor shape for the function is given by $X_i = (H_i, W_i, C_i)$, where W_i and H_i are the spatial
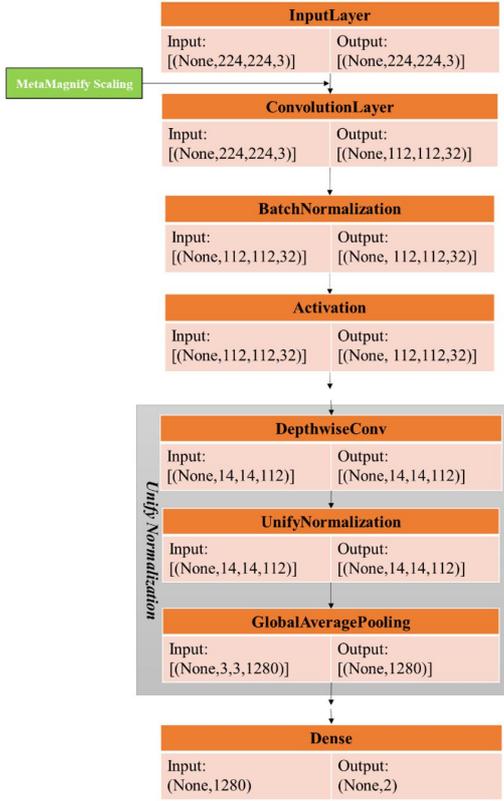
Fig. 4. Model architecture of the proposed MetaMag efficiency model

dimensions. Further, C_i signifies the channel dimension. Moreover, the modified `conv` layer is characterized by a list of composed layers, as shown in equation (4),

$$n = B_k \odot \cdots \odot B_2 \odot B_1(X_1) = \odot_{j=1,\dots,k} B_j(X_j) \,. \tag{4}$$

All layers in each stage of modified layers possess the same convolutional type, while the first layer alone performs the down-sampling method, and the modified `conv` layer is represented in equation (5),

$$n = \odot_{i=1,\dots,m} B_i^{p_i} X_{H_i,W_i,C_i} \,, \tag{5}$$

where $\odot_{i=1,\dots,m} B_i$ is repeated p_i times in stage i, and H_i, W_i, C_i is the shape of input tensor X of layer i. To find the best layer architecture B_i, the model involved MetaMag

scaling that expands the network length p_i, width C_i, and resolution H_i, W_i without altering the predefined B_i in the baseline network. Thus, by fixing the B_i, the MetaMag scaling simplifies the design issues for new resource constraints. However, to improve the accuracy of the proposed model for any resource constraints, an optimization problem is formulated in equation (6),

$$n_{\text{optimum}} = \max_{d,w,r} \text{Accuracy}\left(n(d,w,r)\right), \tag{6}$$

where

$$n(d,w,r) = \odot B_i^{d \cdot L_i}\left(X_{r \cdot H_i, r \cdot W_i, r \cdot C_i}\right),$$

here, (d, w, r) denote the depth, width and resolution of the scaling network, and L_i is the layer at the stage i. Specifically, the modified `conv` layer captures more complex features and gets generalized better in new tasks. But, this network faces difficulty due to vanishing gradient issues. So, the computation is reduced by lowering the training resolution and thus inexpensively fine-tuning at higher resolution. This method is done by implementing the unifying normalization mechanism to normalize activations throughout the network. It combines statistics from LN (Layer Normalization) and BN (Batch Normalization), adapting different batch sizes and model depths. This ensures stable and efficient training across the proposed MetaMag efficiency model. The unified normalization is applied on X, which denotes the unnormalized pre-activations to generate normalized pre-activations $Q_{..c}$ before a nonlinearity Θ and an affine transform finally produce the post-activation function $P_{..c}$, as follows

$$Q_{..c} = \frac{X_{..c} - \mu_c}{\sqrt{\sigma_c^2 + \epsilon}}, \tag{7}$$

$$P_{..c} = \Theta(\gamma_c Q_{..c} + \alpha_c), \tag{8}$$

where c is the index of the channel, $\sqrt{\sigma_c^2}, \mu_c$ denote the standard deviation and mean of X, and α_c, γ_c are the unified normalization's shift parameters and scale in each channel. The ϵ represents the unified normalization's numerical stability constant, and '.' denotes a placeholder for an index. Thus, this foundational principle of unified normalization is significant for successful scaling to deep and large models. Further, the proxy-normalized activation step is applied in equation (8). This step tends to normalize $\Theta(\gamma_c Q_{..c} + \alpha_c)$, where $Q_{..c} \sim N(\alpha_{..c}, (1 + \gamma_c)^2)$ is the proxy variable with variance $(1 + \gamma_c)^2$ and mean $\alpha_{..c}$. These variables are subjected to weight decay to denote that Q is close to normalized. Hence, the unified normalization for each element and the channel is given by equations (9) and (10) (some index placeholders dropped for simplicity),

$$Q_{..b} = \frac{X_{..b} - \mu_b}{\sqrt{\sigma_b^2 + \epsilon}}, \tag{9}$$

$$P_b = \frac{\Theta(\gamma_c Q_{..c} + \alpha_c) - E_{\gamma_c}\left[\Theta(\gamma_c Q_{..c} + \alpha_c)\right]}{\sqrt{\text{Var}_{\gamma_c}[\Theta(\gamma_c Q_{..c} + \alpha_c)] + \epsilon}}, \tag{10}$$

Fig. 5. Examples of original normal images present in the data set.

where b denotes the batch element for the proxy-normalization of P_b; further, $Q_c \sim N(\alpha_c, (1 + \gamma_c)^2), \epsilon$ are numerical stability constants of unified and proxy normalizations, γ_c is the Gaussian proxy variable, and E_{γ_c} represents the measures of central tendency for the variable γ_c. On the other hand, the inclusion of unified normalization at the network leads in a full-batch setting to add the following operations as shown in equation (11),

$$y_{a,c}^l = \frac{y_{a,c}^l - \mu_c(X^l)}{\sigma_c(X^l)}, \quad y_{a,c}^l = \gamma_c^l y_{a,c}^l + \alpha_c^l, \tag{11}$$

where l is the layer, $\sigma_c(X^l)$ and $\mu_c(X^l)$ are the standard deviations and mean of X^l, and α_c^l, γ_c^l denote the shift parameters and channel-wise scale. Finally, the output is driven to the `avg max` pooling layer and then collected from the dense layer.

4. Results and discussion

The effectiveness of the proposed MetaMag efficiency model has been validated by using four different performance measures based on different lower extremity images from the LERA dataset. The experiment was carried out on the Google Colab Notebook Pro version. In total 50 epochs were used in each fold. This section deliberates the results produced by the proposed method in classifying the image samples.

4.1. Exploratory Data Analysis

The Exploratory Data Analysis (EDA) is specifically used to analyze and examine the LERA dataset and thus summarise the main attributes of the dataset. It also visualizes the distribution of data, discovers patterns, locates outliers, and detects correlations. The figure 5 represents the original images present in the LERA dataset.

4.2. Performance measures

The outcome of the proposed system is attained by evaluating the measures: accuracy, precision, recall, specificity, and F1-score. With this output testing accuracy, the improvement of the system is analyzed. Below, TP, TN, FP and FN denote the numbers of true positive, true negative, false positive, and false negative classifications. The probabilities are estimated by the respective relative frequencies.

Accuracy The accuracy is considered as the primary evaluation index in the classification process, which refers to the proportion of input samples that are classified correctly. The accuracy is evaluated as follows

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{12}$$

Precision Precision denotes the probability of the sample that is truly positive among all the samples that are identified to be positive and is given by

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{13}$$

Sensitivity (Also called **recall**; these two names are used interchangeably in the paper, depending on the convention used in the reference sources.) It is the probability of being identified as a positive sample within the actually positive samples. It is denoted as

$$\text{Sensitivity} = \frac{\text{TP}}{\text{FN} + \text{TP}}. \tag{14}$$

Specificity It is the probability of being identified as a negative sample within the actually negative samples. It is denoted as

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \tag{15}$$

F1-score The F1-score is calculated as the harmonic mean of recall and precision and is given by

$$\text{F1-score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}. \tag{16}$$

The above evaluation metrics, or indexes, are used in analyzing the performance of the proposed MetaMag efficiency model.

4.3. Performance analysis

To better verify the efficiency of the proposed model, the obtained results of abnormalities in the body's lower extremities are shown in Figure 6.

From figure 6 it is projected that the proposed model can segment the abnormal part of the image by visualizing it through contrast enhancement. Thus, generalizability was effactually recognized showing a lack of significant decrement in performance.
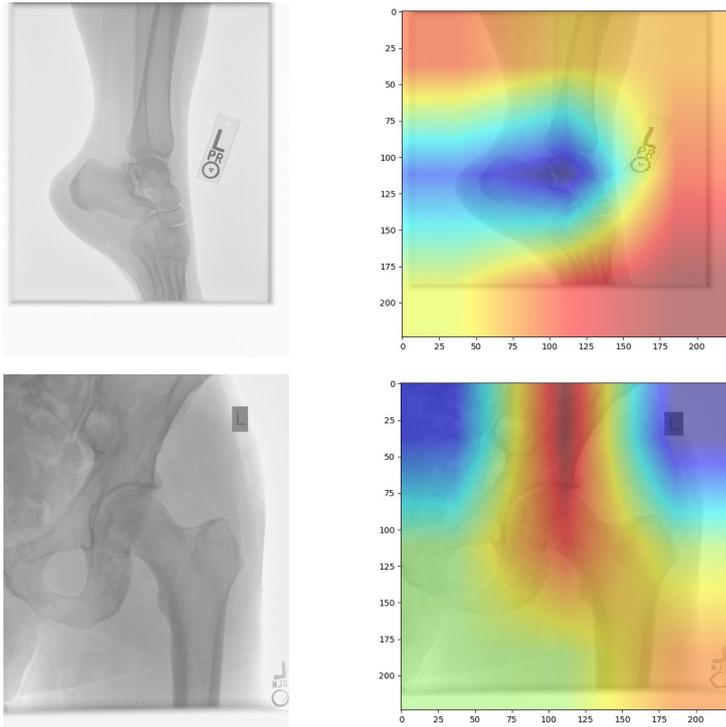
Fig. 6. Abnormalities identified by the proposed MetaMag Efficiency model. Left: original images, right: processed images. In the upper image, the blue color shows high intensity of the abnormality, whereas in the lower image, the red color shows high intensity of the abnormality.

4.4. Internal results

By evaluating the internal test set, the precision, recall, F1-score, and accuracy of the proposed MetaMag Efficiency technique and traditional EfficientNet model are generated. The outcomes are shown in Table 2 and the corresponding graphical representation is displayed in Figure 7.

It is observed that the traditional EfficientNet model produces an accuracy rate of 85%, precision of 94%, recall of 78%, and F1-score of 85%. While, the proposed MetaMag Efficiency model produced an accuracy of 95%, precision of 95%, recall of 97%, and F1-score of 96%. This indicates the improved performance of the proposed method by implementing MetaMag scaling and Unify normalization methods. Figure 8 illustrates the graphical representation of model accuracy and loss.

From figure 8, it can be concluded that the proposed method has produced increased

Tab. 2. Outcome of the proposed and the traditional model.

Model	Precision	Recall	F1-score	Accuracy
Proposed	0.95	0.97	0.96	0.95
EfficientNet	0.94	0.78	0.85	0.85
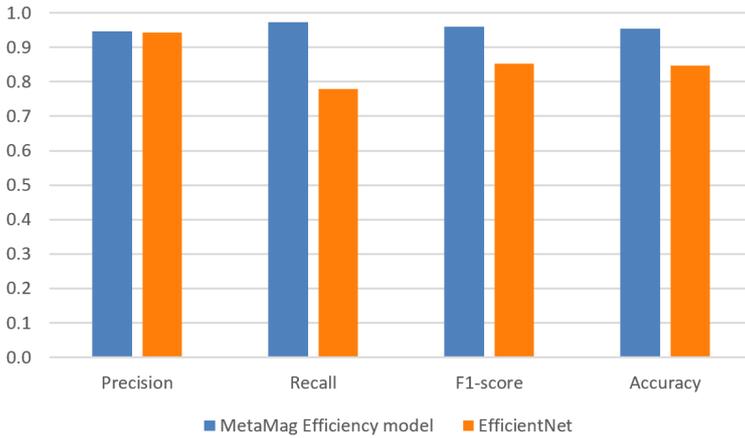
Fig. 7. Graphical representation of performance analysis of the proposed MetaMag Efficiency model and the traditional EfficientNet model.
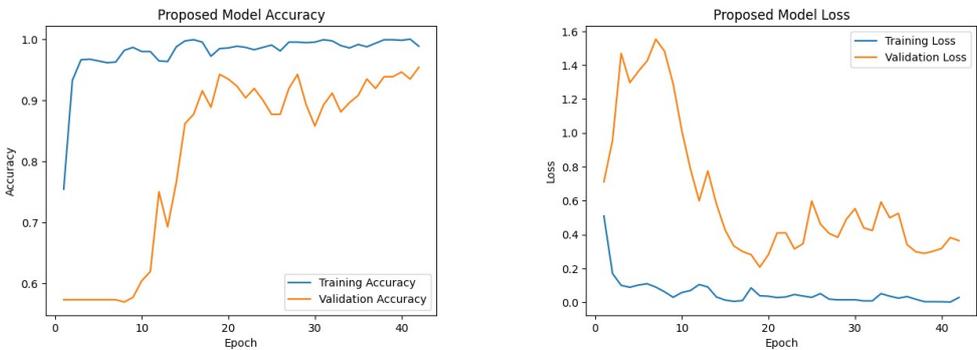
Fig. 8. Accuracy and loss prediction of the proposed MetaMag Efficiency model.
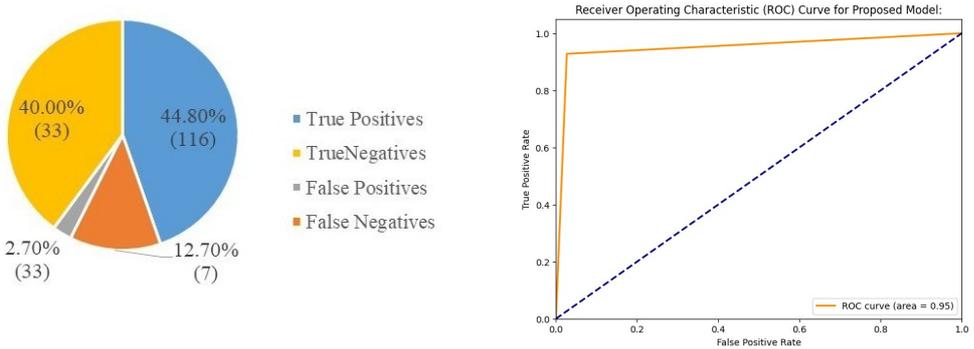
Fig. 9. Data derived from the confusion matrix and ROC for the proposed MetaMag Efficiency model.

accuracy. Both the training curve and validation curve correlate with each other projecting that the train dataset and test dataset are most probably similar to each other. Further, the data derived from the confusion matrix are drawn for the proposed method to analyze effectiveness. The model loss indicates how the model's prediction was on the input samples. If the loss is minimal, then the efficacy of the proposed approach will be enhanced. In Figure 8, the x-axis denotes the loss and the y-axis signifies the number of model training epochs. It is noted that the validation accuracy is higher than the training accuracy for some epochs. Both the training and validation curve follows a uniformity as the number of epochs increases. This denotes that the loss decreased with an increase in accuracy. The data derived from the confusion matrix and the ROC of the proposed model are represented in Figure 9.

The confusion matrix, also known as the error matrix, represents the counts from predicted and actual values. The True Positives value represents the number of positive samples that are accurately classified, while True Negatives denotes the number of negative samples categorized correctly. False Positives value signifies the number of actual negative samples classified as positive, and False Negatives is the number of actual positive samples classified as negative. From Figure 9 it is inferred that 144 samples were correctly classified as normal images, and 104 abnormal samples were classified accurately. Only 4 normal samples were misclassified as abnormal, and 8 abnormal samples as normal. With minimum error, the precision of the proposed approach is improved. Additionally, the area under the ROC curve of the proposed model is found to be 0.93, indicating improved performance. Further, the performance of the conventional EfficientNet is also analyzed. The data derived from the confusion matrix and ROC of the traditional EfficientNet model are shown in Figure 10.

The data derived from the confusion matrix of the traditional EfficientNet model is analyzed, and it is found that 116 normal samples are correctly classified, and 104
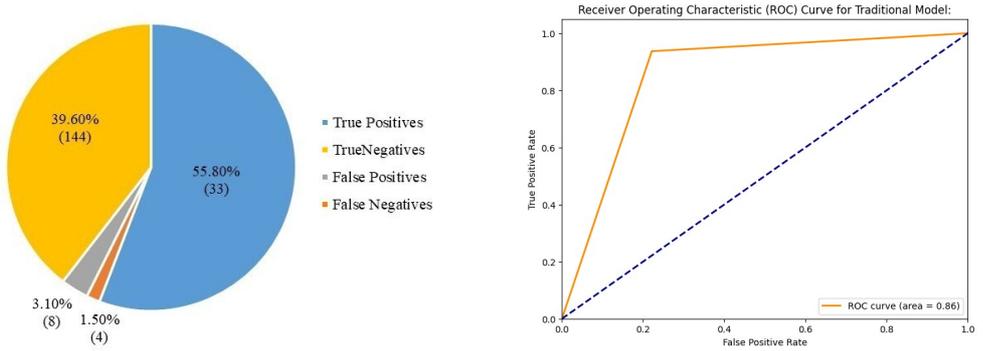
Fig. 10. Data derived from the confusion matrix and ROC for the traditional EfficientNet model.
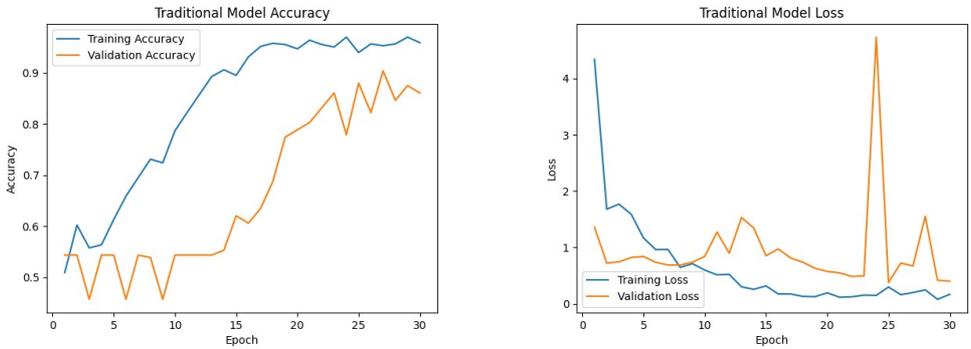
Fig. 11. Accuracy and loss prediction of the traditional EfficientNet Model.

abnormal images are classified accurately, while 33 normal images are wrongly classified as abnormal and 7 abnormal images are classified incorrectly as normal samples. The count of correctly classified samples is less than the count classified by the proposed model. Further, the area under the ROC curve of the traditional model is 0.86, denoting decreased accuracy and performance. Then, the model accuracy and loss prediction for the conventional method is shown in Figure 11.

The training and testing curves are partially correlated in the model accuracy plot, denoting decreased accuracy. Further, the loss plot denotes that both the loss curve interlinks with each other, representing increased model loss. This denotes that increased model loss leads to reduced performance of the model.

Tab. 3. Comparison of the performance of the proposed model and the existing model from [18].

Model	Recall	Specificity	Accuracy
DCNN Triquetral fracture (n=50) – 2-stage	0.96	0.88	0.92
DCNN Triquetral fracture (n=50) – 1-stage	0.96	0.64	0.80
Second fracture (n=24) – 2-stage	0.917	0.917	0.917
Second fracture (n=24) – 1-stage	0.917	0.917	0.917
Proposed	**0.97**	**0.93**	**0.95**

4.5. Comparative analysis

The comparison of the proposed method with other existing methods enumerates the efficiency of the proposed system. Here, the study compares the existing DenseNet-161 model in terms of lower extremities' accuracy, sensitivity, and specificity. The outcome of the conventional model and the proposed system is exemplified in Table 3.

Table 3 indicates that the proposed model attained better values than the existing models. It attained 95% of sensitivity, 97% of specificity and 95% of accuracy which shows the value of the proposed efficient model. Table 4 depicts the comparative analysis of the proposed and another existing model.

From Table 4 and Figure 12 it can be inferred that the existing DenseNet-161 model produced an accuracy of 79%, precision of 97%, and recall of 66%. Whereas the proposed MetaMag Efficiency model produced importantly improved overall accuracy of 95%, precision of 95% – slightly worse, and recall of 97% – improved.

Only a few studies have focused on detecting bone fractures in DM patients. So, only a limited comparison is provided to analyze the model's working. Thus, from analyzing using different evaluation indicators, it is identified that the proposed model has achieved improved performance compared to other existing models. The basic EfficientNet model tends to provide limited performance, whereas the proposed MetaMag efficiency model provides improved performance due to the implementation of the MetaMag scaling and Unify normalisation. The MetaMag scaling supports the model in enlarging the radiographic image and finely detecting significant patterns of abnormalities in the bone. Further, the unified normalization reduces the losses produced by the input samples and thus increases the model's efficiency.

Tab. 4. Comparison of the performance of the proposed model and the DenseNet model [30].

Model	Precision	Recall	Accuracy
DenseNet-161 [30]	0.97	0.66	0.79
Proposed	0.95	**0.97**	**0.95**
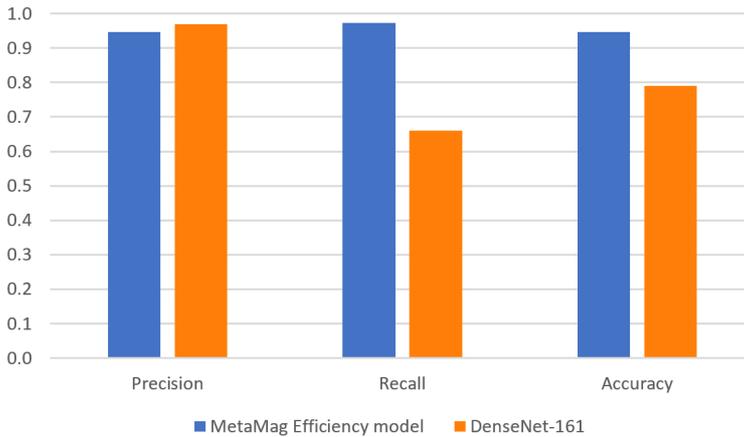
Fig. 12. Graphical representation of the comparison of the proposed model and the DenseNet-161 model [30].

5. Conclusion

Various deep learning methods are involved in diagnosing various diseases and have produced efficient outcomes. In that case, the previous phase concentrated on detecting foot ulcers in diabetes mellitus (DM) patients by using the Deep Convolutional Neural Network (DCNN) based Xception model. This approach produced improved outcomes and aided in efficiently classifying healthy and diabetic foot ulcer (DFU) images. On the other hand, the present phase focused on identifying fractures in diabetes patients. DM is associated with several other factors, and delay in treatment may lead to complex patient risks. Once a fracture occurs in diabetes patients, it is difficult to cure, and abnormalities exploit the routine lifestyle of patients. So, early detection of fractures can help physiologists efficiently cure the complications. Hence, the proposed approach implemented a MetaMag efficiency model to detect and classify normal and abnormal images from the given input radiograph images. Along with the classifier, MetaMag scaling and Unify normalization approaches were used to effectively obtain the fine details of input samples and reduce the loss that occurred in the proposed system. The outcomes of the proposed method produced an accuracy of 95%, compared with the traditional EfficientNet model, which produced an accuracy of 85%. This denoted the improved performance of the proposed MetaMag efficiency model. The study can be further improved by using different approaches of deep learning algorithms to produce higher classification accuracy.

References

[1] R. Ali, J. H. Chuah, M. S. A. Talip, N. Mokhtar, and M. A. Shoaib. Structural crack detection using deep convolutional neural networks. *Automation in Construction*, 133:103989, 2022. doi:10.1016/j.autcon.2021.103989.

[2] M. J. Awan, M. S. M. Rahim, N. Salim, M. A. Mohammed, B. Garcia-Zapirain, et al. Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach. *Diagnostics*, 11(1):105, 2021. doi:10.3390/diagnostics11010105.

[3] R. Bagaria, S. Wadhwani, and A. K. Wadhwani. Bone fracture detection in X-ray images using convolutional neural network. In: *SCRS Conference Proceedings on Intelligent Systems*, pp. 459–466. SCRS, India, 2022. doi:10.52458/978-93-91842-08-6-43.

[4] Z. Cao, L. Xu, D. Z. Chen, H. Gao, and J. Wu. A robust shape-aware rib fracture detection and segmentation framework with contrastive learning. *IEEE Transactions on Multimedia*, 25:1584–1591, 2023. doi:10.1109/TMM.2023.3263074.

[5] W. Chen, D. HolcDorf, M. W. McCusker, F. Gaillard, and P. D. Howe. Perceptual training to improve hip fracture identification in conventional radiographs. *PloS One*, 12(12):e0189192, 2017. doi:10.1371/journal.pone.0189192.

[6] P. Chłąd and M. R. Ogiela. Deep learning and cloud-based computation for cervical spine fracture detection system. *Electronics*, 12(9):2056, 2023. doi:10.3390/electronics12092056.

[7] M. Davenport and M. P. Oczypok. Knee and leg injuries. *Emergency Medicine Clinics*, 38(1):143–165, 2020. doi:10.1016/j.emc.2019.09.012.

[8] M. R. Delavar. Hybrid machine learning approaches for classification and detection of fractures in carbonate reservoir. *Journal of Petroleum Science and Engineering*, 208:109327, 2022. doi:10.1016/j.petrol.2021.109327.

[9] C. Germann, A. N. Meyer, M. Staib, R. Sutter, and B. Fritz. Performance of a deep convolutional neural network for MRI-based vertebral body measurements and insufficiency fracture detection. *European Radiology*, 33(5):3188–3199, 2023. doi:10.1007/s00330-022-09354-6.

[10] N. Gougoulias, H. Oshba, A. Dimitroulias, A. Sakellariou, and A. Wee. Ankle fractures in diabetic patients. *EFORT Open Reviews*, 5(8):457–463, 2020. doi:10.1302/2058-5241.5.200025.

[11] O. Q. Groot, M. E. R. Bongers, P. T. Ogink, J. T. Senders, A. V. Karhade, et al. Does artificial intelligence outperform natural intelligence in interpreting musculoskeletal radiological studies? A systematic review. *Clinical Orthopaedics and Related Research*, 478(12):2751, 2020. doi:10.1097/CORR.0000000000001360.

[12] B. Guan, J. Yao, G. Zhang, and X. Wang. Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network. *Pattern Recognition Letters*, 125:521–526, 2019. doi:10.1016/j.patrec.2019.06.015.

[13] K. Karthik and S. S. Kamath. MSDNet: A deep neural ensemble model for abnormality detection and classification of plain radiographs. *Journal of Ambient Intelligence and Humanized Computing*, 14:16099–16113, 2023. doi:10.1007/s12652-022-03835-8.

[14] C. Kokkotis, S. Moustakidis, G. Giakas, and D. Tsaopoulos. Identification of risk factors and machine learning-based prediction models for knee osteoarthritis patients. *Applied Sciences*, 10(19):6797, 2020. doi:10.3390/app10196797.

[15] S. Kumar, P. Goswami, and S. Batra. Fuzzy rank-based ensemble model for accurate diagnosis of osteoporosis in knee radiographs. *International Journal of Advanced Computer Science and Applications*, 14(4):262–270, 2023. doi:10.14569/IJACSA.2023.0140430.

[16] Y. Ma and Y. Luo. Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network. *Informatics in Medicine Unlocked*, 22:100452, 2021. doi:10.1016/j.imu.2020.100452.

[17] Stanford University School of Medicine. LERA- Lower Extremity RAdiographs, 2024. https://aimi.stanford.edu/lera-lower-extremity-radiographs, [Accessed: September 19, 2024].

[18] M. Ren and P. H. J. S. R. Yi. Deep learning detection of subtle fractures using staged algorithms to mimic radiologist search pattern. *Skeletal Radiology*, 51(2):345–353, 2022. doi:10.1007/s00256-021-03739-2.

[19] A. Sasidhar, M. Thanabal, and P. Ramya. Efficient transfer learning model for humerus bone fracture detection. *Annals of the Romanian Society for Cell Biology*, 25(2):3932–3942, 2021. http://annalsofrscb.ro/index.php/journal/article/view/1398.

[20] T. Schmidt, N. M. Simske, M. A. Audet, A. Benedick, C.-Y. Kim, et al. Effects of diabetes mellitus on functional outcomes and complications after torsional ankle fracture. *Journal of the American Academy of Orthopaedic Surgeons*, 28(16):661–670, 2020. doi:10.5435/JAAOS-D-19-00545.

[21] H. Sun, X. Wang, Z. Li, A. Liu, S. Xu, et al. Automated rib fracture detection on chest X-ray using contrastive learning. *Journal of Digital Imaging*, 36(5):2138–2147, 2023. doi:10.1007/s10278-023-00868-z.

[22] Y. L. Thian, Y. Li, P. Jagmohan, D. Sia, V. E. Y. Chan, et al. Convolutional neural networks for automated fracture detection and localization on wrist radiographs. *Radiology: Artificial Intelligence*, 1(1):e180001, 2019. doi:10.1148/ryai.2019180001.

[23] K. A. Thomas, Ł. Kidziński, E. Halilaj, S. L. Fleming, G. R. Venkataraman, et al. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. *Radiology: Artificial Intelligence*, 2(2):e190065, 2020. doi:10.1148/ryai.2020190065.

[24] M. Tian, B. Li, H. Xu, D. Yan, Y. Gao, et al. Deep learning assisted well log inversion for fracture identification. *Geophysical Prospecting*, 69(2):419–433, 2021. doi:10.1111/1365-2478.13054.

[25] M. Varma, M. Lu, R. Gardner, J. Dunnmon, N. Khandwala, et al. Automated abnormality detection in lower extremity radiographs using deep learning. *Nature Machine Intelligence*, 1(12):578–583, 2019. doi:10.1038/s42256-019-0126-0.

[26] S. Verma, S. Kulshrestha, C. Rajput, and S. Patel. Detecting bone fracture using transfer learning. In: O. P. Verma, S. Roy, S. C. Pandey, and M. Mittal, eds., *Advancement of Machine Intelligence in Interactive Medical Image Analysis*, pp. 215–228, Algorithms for Intelligent Systems. Springer Singapore, 2020. doi:10.1007/978-981-15-1100-4_10.

[27] M. Wu, Z. Chai, G. Qian, H. Lin, Q. Wang, et al. Development and evaluation of a deep learning algorithm for rib segmentation and fracture detection from multicenter chest CT images. *Radiology: Artificial Intelligence*, 3(5):e200248, 2021. doi:10.1148/ryai.2021200248.

[28] L. Xue, W. Yan, P. Luo, X. Zhang, T. Chaikovska, et al. Detection and localization of hand fractures based on GA_Faster R-CNN. *Alexandria Engineering Journal*, 60(5):4555–4562, 2021. doi:10.1016/j.aej.2021.03.005.

[29] D. P. Yadav, A. Sharma, S. Athithan, A. Bhola, B. Sharma, et al. Hybrid SFNet model for bone fracture detection and classification using ML/DL. *Sensors*, 22(15):5823, 2022. doi:10.3390/s22155823.

[30] J. Zhang, Z. Li, S. Yan, H. Cao, J. Liu, et al. An algorithm for automatic rib fracture recognition combined with nnU-Net and DenseNet. *Evidence-Based Complementary and Alternative Medicine*, 2022(1):5841451, 2022. doi:10.1155/2022/5841451.

[31] K. Üreten, H. F. Sevinç, U. İğdeli, A. Onay, and Y. Maraş. Use of deep learning methods for hand fracture detection from plain hand radiographs. *Turkish Journal of Trauma and Emergency Surgery*, 28(2):196, 2022. doi:10.14744/tjtes.2020.06944.

S. Rajeashwari is a Full Time Research Scholar in the Department of Computer Science, Sri S. Ramasamy Naidu Memorial College, Affiliated to Madurai Kamaraj University, Madurai, India. Her major research areas include Data Mining, Knowledge Engineering, Image Processing, Medical Image Analysis, and Machine Learning.

Dr. K. Arunesh is an Associate Professor in the Department of Computer science working in S. Ramasamy Naidu Memorial College, Affiliated to Madurai Kamaraj University, Madurai, India. He has 34 years of teaching experience. His major research areas include Machine Learning, Web Usage Mining, Recommender System, Data Mining, and Computational Intelligence. He has over 40 publications in refereed journals and serves as a reviewer for several esteemed journals. He has also been an advisory committe member for various conferences.