JSR Publications

Available Online

J. Sci. Res. **17** (3), 789-807 (2025)

JOURNAL OF SCIENTIFIC RESEARCH

www.banglajol.info/index.php/JSR

An Ensemble Methodology for Osteoporosis Prediction Using Knee X-Ray Images

P. Kour, S. Kumar, S. Shastri*, V. Mansotra

Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India

Received 9 January 2025, accepted in final revised form 28 May 2025

Abstract

Osteoporosis is the most prominent chronic bone disorder characterized by a deficit bone mineral density (BMD), which elevates the probability of bone fractures. An early and precise osteoporosis diagnosis increases the patient's survival rate. X-ray imaging is the most affordable and accessible method for diagnosing bone diseases. Still, manually interpreting X-rays for osteoporosis is laborious and time-consuming, and choosing high-performance classifiers is a highly challenging task. To address the above issues, this paper presents an ensemble-based model that uses knee X-ray images to predict osteoporosis as a binary class (normal and osteoporosis). The model used different transfer learning techniques and a custom CNN to detect osteoporosis in this work. It is found that the ensemble model has produced exceptional outcomes, particularly in terms of accuracy for binary classification. Additionally, the model's efficiency was tested and compared with other deep learning models, which indicated that the ensemble model is more robust than recent DL approaches. Hence, our methodology may save surgeons time while simultaneously enhancing patient outcomes.

Keywords: Knee X-ray images; Osteoporosis; Deep Learning; Transfer learning; Ensemble learning.

doi: https://dx.doi.org/10.3329/jsr.v17i3.79003 J. Sci. Res. **17** (3), 789-807 (2025)

1. Introduction

Osteoporosis is an asymptomatic and multifactorial bone disorder characterized by a reduction in bone density or bone quality and micro-architectural bone tissue deterioration, subsequently enhancing bone fragility and elevating the risk of bone fracture. It is considered a silent disease as there are often no symptoms before the first fracture, and it is seen in all age groups, genders, and races, especially postmenopausal women who are at the highest risk. Over 200 million individuals worldwide, one in three women and one in five men over 50, are expected to have osteoporotic fractures, according to estimates from the International Osteoporosis Foundation [1]. Osteoporosis not only induces fractures but also causes people to become bedridden, which can lead to secondary complications that can be fatal over time. Osteoporosis is difficult to identify in its early stages because of a lack of

* *Corresponding author*: sourabhshastri@gmail.com

recognizable symptoms and can result in patient death in many circumstances. It is a widespread bone disorder that causes millions of fractures and poses a global medical risk. Early identification and treatment of osteoporosis can reduce death rates while increasing survival rates.

Osteoporosis-related fractures impose a heavy burden on both individuals and society, as their treatment costs a huge amount of budget for economies. It accelerates the death rate, morbidity, and disability while significantly reducing quality of life. It causes fractures of the humerus, pelvis, thoracolumbar vertebrae, distal forearm, hip, etc. Females are at a higher risk of tibial and fibular fractures, and as the population ages, osteoporotic fractures surrounding the knee have become more common [2]. The X-ray image illustrates that due to decreased bone density, the osteoporotic knee has more gray-level intensity than normal. An annual death rate of 22 % is found in elderly patients who endure femoral fractures, with a poor quality of life, and approximately fifty percent of all knee fractures occur in older individuals over 50 [3]. Some risk factors for osteoporosis are modifiable, such as diet, weight reduction, air pollution, stress, and lifestyle variables, while some factors are non-modifiable, such as older age, gender, past fractures, and reproductive factors [4]. Furthermore, numerous factors, such as cancer therapies, bone cancer, or certain cancer types that metastasize to the bones, can contribute to the development of osteoporosis [5]. Diagnostic imaging modalities used for osteoporosis detection include digital X-ray radiogrammetry (DXR), quantitative ultrasound (QUS), magnetic resonance imaging (MRI), quantitative computed tomography (QCT), high-resolution peripheral quantitative computed tomography (HR-pQCT), and radiography. The Dual Energy X-ray Absorptiometry Technique (DXA) is the method most commonly used in healthcare to diagnose osteoporosis [6]. An X-ray is the traditional and widely used screening technique for acquiring images of almost every single bone in the body, including the bones of the hand, ankle, elbow, pelvic region, and vertebrae. Furthermore, these methods are expensive, have a low spatial resolution, and have high radiation doses, making them unsuitable for screening and unable to identify osteoporosis until the disease has progressed. DXA is the preferred norm for determining bone mineral density (BMD), however, it only captures information concerning bone strength and ignores the contribution of clinical risk factors and bone constraints (trabecular bone geometry and score) [7,8]. Furthermore, human diagnostics is time-consuming, prone to errors, and does not fulfill social expectations. Considering these limitations, a cost-effective, efficient, and novel approach for automatically detecting malignancy is required. As a result, the researchers develop computer-aided diagnostic tools by using recent developments in imaging technology to evaluate medical images through computer algorithms.

Computer-aided detection and diagnosis are being utilized to improve the identification or diagnosis of osteoporosis. Medical, pharmacological, and fundamental biology have benefited from the advent of AI technologies, which have improved performance in these domains, and human specialist-level performance has been attained [9]. Furthermore, smarter diagnostic tools have been invented to solve diagnostic problems and improve efficiency in healthcare [10]. Deep learning, a subfield of artificial intelligence and high-

level neural networks that resembles the human brain, solves complex problems, facilitates automatic feature extraction, and is rapidly being applied in both fundamental and clinical cancer research and bone-associated disorders [11]. Among other DL approaches, Convolutional Neural Networks (CNNs) are the most well-known and are gaining more attention from researchers. CNN is a state-of-the-art technique for processing massive medical image data sets to determine and test new imaging attributes [12-15]. Numerous CNN research studies have embraced transfer learning strategies, which reuse the knowledge learned from a network to enhance performance by freezing initial layers and fine-tuning the training parameters in the final layers [16,17]. It eliminates the necessity for large training datasets while simultaneously performing exceptionally well in tumor detection, microbleed segmentation, tumor segmentation, and many other areas [18]. Although the researchers have proposed methodologies for osteoporosis detection at several bone sites, such as the spine, upper extremity, hip region, and tooth, there hasn't been much study done on transfer learning techniques to identify osteoporosis in the knees [19].

The automated categorization of biological samples has received a great deal of interest in recent years. Some of the previous research in the field of osteoporosis detection and classification using deep learning and transfer learning has been discussed in the literature. Several machine learning techniques were utilized to build a prediction model for femoral neck osteoporosis [20]: Yang *et al.* predicted the presence of osteoporosis [21]: Dadsetan *et al.* used radiomics on clinically accessible X-ray images for predicting the risk of osteopenia and osteoporosis [22]. Twenty machine-learning techniques were assessed based on four diagnostic factors (age, sex, height, and weight) for osteoporosis risk prediction and categorized into osteoporosis and non-osteoporosis [23]. A novel vibroacoustic approach was presented to detect osteoporosis from the tibia's impulse responses in vivo along with an artificial neural network, and the obtained results showed that it could be suitable for screening tests for the early detection of osteoporosis [24]. Lim *et al.* assessed the predictive ability of machine learning analysis with radionics features and abdominal-pelvic CT (APCT) to diagnose femoral osteoporosis [25]. Based on clinical data and CT images, Liu *et al.* developed a three-layered hierarchical model by utilizing six different machine learning techniques to identify osteoporosis [26].

A novel approach has been developed to predict whether the patient is osteoporotic or healthy from the radiographic image of the bone [27]. The authors examined the use of deep learning to identify osteoporosis using dental panoramic radiographs [28]. Furthermore, an ensemble model-based osteoporosis classifier was developed, to which patient clinical variables were incorporated. The outcomes showed that osteoporosis could be properly classified using deep learning with CNN, and the performance could be enhanced using an ensemble model incorporating patient covariates. In another research study, four distinct transfer learning models with fine-tuning parameters were utilized to distinguish osteoporosis using dental panoramic radiographs (DPRs), labeled based on T-score [29]. Tassoker *et al*. presented several deep learning methods for predicting radiological variations that resemble osteoporosis on panoramic radiographs in women above the age of fifty and validated them with the 5-fold cross-validation approach [30]. To optimize ELM

to differentiate between osteoporotic and normal data, an HMBA-ELM classifier was developed [31]. The algorithm was developed by merging an artificial algae algorithm, a multi-light source algorithm (AAAML), and a Monarch butterfly optimization (MBO) algorithm, and the evolutionary phase of AAAML was called Evolution-migrated AAAML. The methodology was assessed using 10 independent runs of 10-fold cross-validation on three distinct osteoporosis datasets. It was found that the proposed HMBA-ELM successfully classified osteoporotic datasets with outstanding outcomes.

The outcomes of a computer-aided diagnosis (CAD) system based on deep convolutional neural networks (DCNNs) to identify osteoporosis utilizing radiographs were compared with the diagnosis made by oral and maxillofacial radiologists by Lee *et al.* [32]. The convolutional neural network was trained, validated, and tested using both single-column and multicolumn deep convolutional neural networks. To facilitate the early and trustworthy diagnosis, Gaudin *et al.* showed that a deep learning algorithm could accurately identify osteoporosis using panoramic radiographs [33]. A deep neural network (DNN) model and a nonlocal neural network (NLNN) model based on VGG16 were integrated to diagnose osteoporosis from hip radiography [34]. The outcomes of the study showed that the model was considered to be among the best and most effective screening tools for osteoporosis prediction. To assess patient X-ray images of the chest and detect COVID-19-positive people, Kumar and Rani introduced the "LiteCovidNet" model, which is based on deep neural networks [35]. The COVID-19 case was divided into two classes by the model: a multi-class and a binary class. An optimal artificial neural network was proposed by Umar *et al.* to forecast the closing price for each day of the National Stock Exchange's NIFTY-50 index [36].

According to the literature discussed above, DL-based diagnostic models outperform machine learning approaches for osteoporosis diagnosis or classification. We have also concluded that some models' performance is unsatisfactory. Many methods for detecting and classifying bone fractures have been proposed in the past, but less work has been done for predicting and classifying knee osteoporosis with X-ray. This fact motivates the researchers to concentrate on developing an effective model for osteoporosis prediction. In this study, a methodology has been developed by ensembling transfer learning and a custom CNN model that can efficiently predict knee osteoporosis in patients using X-ray images. The proposed ensemble model has been evaluated based on accuracy and specificity, evaluating measures.

The following are the main objectives of this study:

- To demonstrate that transfer learning-based algorithms can effectively diagnose osteoporosis using a publicly available X-ray images dataset.
- Investigate a suitable CNN framework for effective prediction and uncover potential hints that aid performance.
- The proposed Ensemble approach distinguishes between normal and osteoporosis as a binary classification.

- Explore how transfer learning methods and custom CNNs work together. It is a one-of-a-kind, significant, practical, and successful ensemble methodology for predicting knee X-ray images that yields positive results.
- Compare and evaluate the outcomes of the proposed methodology with those of existing approaches.

The format of the manuscript is as follows: The osteoporosis imaging analysis procedure is thoroughly explained in the Materials and Methods section. Furthermore, it discusses the phases involving data preparation and classification utilizing deep learning and machine learning techniques. Using a variety of performance metrics, the Results section analyses feature categories and assesses the outcomes of deep learning methods. The Discussion section provides an overview of the study's overall achievements, lays out the context for the results, and offers suggestions for future research possibilities.

2. Materials and Methods

The section provides a detailed description of the several processes that the study undertook, starting with data processing and continuing with learning and prediction utilizing ensemble models and deep learning.

2.1. *Dataset description*

This section has covered the dataset that is being utilized in the experiments. For the prediction of osteoporosis, an X-ray image dataset obtained from a source has been utilized, as discussed below.

The dataset has been taken from Kaggle, open-source, and community-maintained platforms. Table 1 includes a link to where the data for the experiment was collected. X-ray images of individuals with normal and osteoporosis patients are included in the dataset. The dataset has two classes (binary classification). Fig. 1 depicts the data distribution as a bar chart for binary classifications (normal and osteoporosis). Table 1 displays the cases of both normal and osteoporosis patients, as well as their references listed in the next row. Out of the 360 total X-ray images, 180 are of a healthy individual and 180 are of a patient with osteoporosis. There were 360 X-ray images utilized in the experiment, with the total number of images included by a certain class being shown in the number of images column.

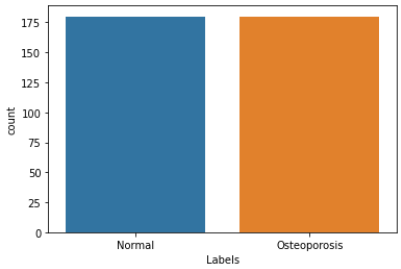

Fig. 1. Dataset distribution.

Table 1. Dataset description.

Classes	Number of Images	Reference
Osteoporosis	180	[37]
Normal	180	

Total X-Ray Images used for the experiment: 360

2.2. *Methodology*

In this section, the process of the ensemble model for the prediction of osteoporosis is described. Additionally, there has been a discussion on the utilized transfer learning models, the custom CNN architecture, the proposed architecture's working, and the algorithm.

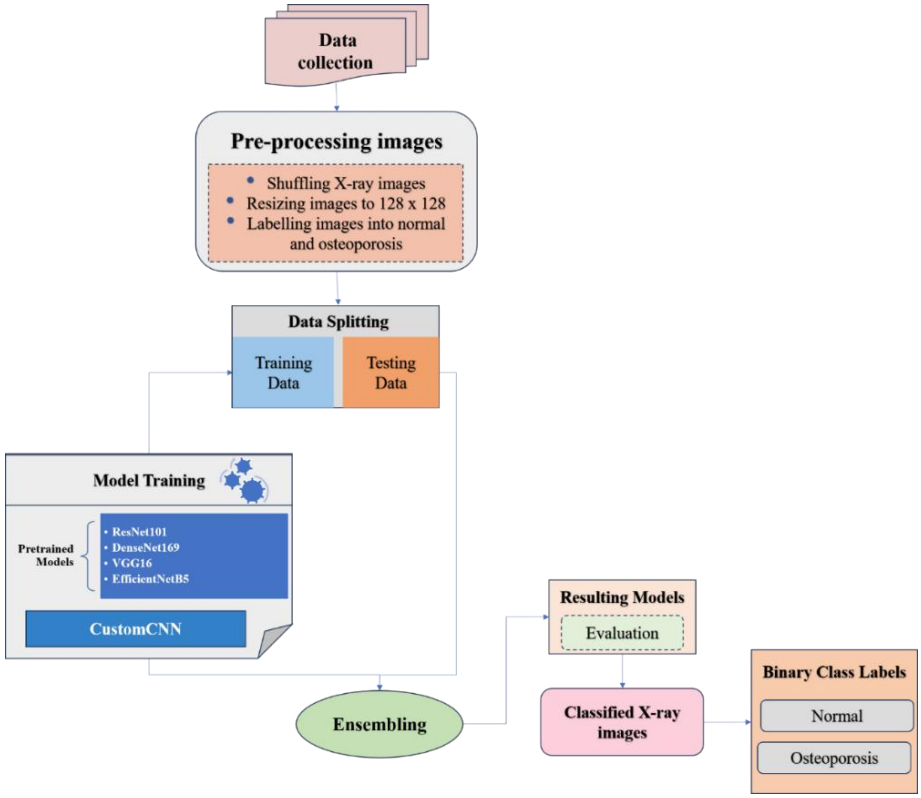

Fig. 2. Workflow for ensemble methodology.

Initially, the knee X-ray image dataset was extracted from the publicly accessible Kaggle platform. The data set needs to be enhanced before model training to improve the model's robustness as well as avoid excessive attention to meaningless features. The X-ray images in the dataset have been preprocessed, which includes resizing, shuffling, and labeling into two classes: normal and osteoporosis. There are differences in the resolution of the images within the datasets. As a result, images have been resized to 128x128 pixels. The

preprocessed data was divided into training -80 % (D_{Xtt}) and testing – 20 % (D_{Xts}). The training data has been employed to construct the structure of the ensemble model. The model has been trained over 200 epochs using a batch size of 25 to avoid the overfitting problem. After configuring the hyperparameters, the pretrained models (ResNet101, DenseNet169, VGG16, and EfficientNetB5) and the custom CNN model have been trained on the training dataset. Custom CNN (WXC) and pretrained models (WXN) weights have been saved and utilized for ensemble training. The resulting model has been obtained from the ensembling model, and the weights will be saved. However, testing data has been employed in the model's evaluation. Based on the results, the knee X-ray images have been categorized as osteoporosis and normal. Fig. 2 depicts the workflow block diagram for osteoporosis prediction using X-ray images of the knee, and the algorithm is illustrated in Algorithm 1.

Algorithm 1: Osteoporosis Prediction

1 Input: Knee X-ray Images Dataset
2 Dataset Preparation: Shuffling, Resizing and Labelling the image dataset
Splitting the dataset
D_X:{ Training data D_{Xtt} and Testing data D_{Xts}}

3 Initialize and varying hyperparameters
4 Pretrained models are built up as follows:
Vgg16, DenseNet169, ResNet101, EfficientNetB5, custom CNN model setup
5 Training:
a) For all pretrained models do
For all epochs do
Execute the pre-trained models that have been defined.
Hyperparameters have been modified to conserve weights (W_{XN}) and choose the optimal pretrained model.
End for
End for
b) For all epochs do
Train the custom CNN model
Hyperparameters are adjusted to select the optimal model and save weights (W_{XC})
End for
c) For all epochs do
Train the ensemble model using save weights W_{XN} and W_{XC}
End for
6 Testing
For D_{Xt} do
Test the pretrained models and custom CNN
Test ensemble model
Classify into binary class and generate the confusion matrix
Calculate Accuracy and Specificity
End for
7 Output: Binary class labels (Normal and Osteoporosis)

2.3. *Convolutional Neural Network*

CNN is a unique type of deep neural network with intermediate layers built on the convolution idea and has demonstrated remarkable performance in image identification, segmentation, and classification [38]. The fundamental components of CNNs include of input layer, convolutional layers, pooling layers, fully connected layers, and an output layer [39]. It has several architectures, including AlexNet, VGG, Inception, NasNet, and DenseNet, based on how effectively they can train [40]. In the study, four pretrained models and a CNN model have been utilized for prediction by fine-tuning their parameters.

2.4. *Transfer learning models*

Transfer Learning is an approach that allows you to train a new model with minimum fine-tuning or training parameters by applying previously learned model knowledge [41]. It reduces training time, trains with limited data, and enhances neural network performance [42]. There are various transfer learning models available, including Inception, ResNet, EfficientNet, Vgg, and many more. In the study, four pretrained models such as ResNet101, DenseNet169, Vgg16 and EfficientNetB5 has applied and described in Table 2.

Table 2. Description of pretrained models.

Architecture	Trained Layers	Description
ResNet101	101	It has 25.6 M parameters and a residual block stack with three layers. Residual blocks employ skip connections, which allow the system to train faster by transmitting the activation to a layer deeper in the network and address the problem of vanishing gradients.
DenseNet169	169	It has a total of 14.3M parameters and solves the vanishing gradient problem, has a strong feature propagation approach, reduces trainable parameters, and encourages the reuse of features [43,44].
VGG16	16	The Visual Geometry Group-16 is constructed with 13 Conv2D blocks and three fully interconnected layers [45]. It has 138.4 M parameters and employs stacks of convolutional layers with smaller kernels of size 3x3.
EfficientNetB5	312	The model's layers were trained at 456 x 456 resolution utilizing ImageNet-1k and have 30.6M parameters.

2.5. *Custom CNN architecture*

The section goes into depth on the custom CNN architecture. The custom CNN model comprises two phases, which are training and testing. During the training phase, feature extraction has been performed. After training, the custom CNN developed has been used for the classification of osteoporosis. Following the initial extraction of the feature, classification has been performed and the OsteoporosisNet architecture has been trained. During the testing phase, the proposed framework has been assessed. Finally, the model

evaluation findings have been used to classify the X-ray images into two categories: normal or osteoporosis. When this process is complete, the model weights are saved for ensemble processing. Fig. 3 depicts the workflow of the custom CNN framework for the prediction of osteoporosis.

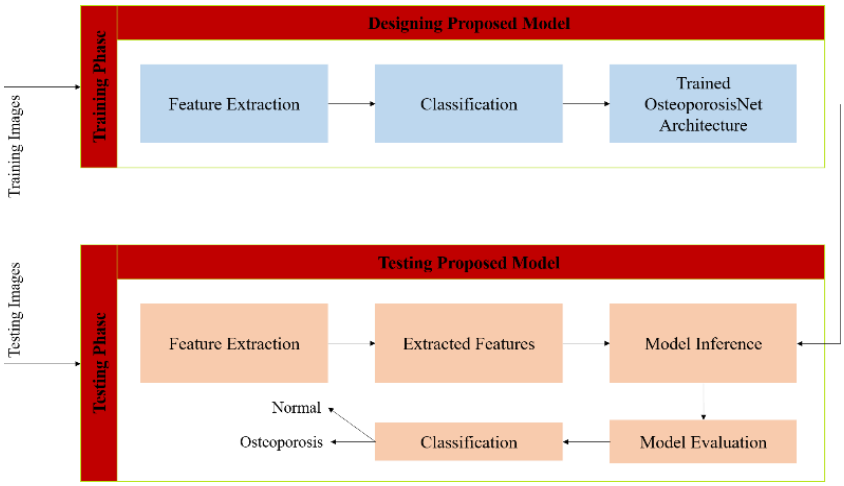

Fig. 3. Flow diagram of Custom CNN work.

The architecture of the custom CNN is shown in Fig. 4, whereas Table 3 describes the custom CNN layers. Table 3 describes the custom CNN layers. This model represents a deep learning model for predicting osteoporosis in which it transforms raw pixels of the input image using convolutional or feature extraction layers, batch normalization layers, leaky RELU layers, and pooling layers. Finally, the last fully connected layer, which receives the feeds, uses class scores or probabilities to classify the input into the class with the highest likelihood.

The first layer, through which data enters the network, is called the input layer. Here, the input layer takes knee X-ray images as input. The task of extracting features from the image has been assigned to the feature extraction layer, which consists of four convolutional layers with varying filters (32, 64, 256, and 512) and kernel sizes of 2 × 2 and 3 × 3. The integration of an activation function aids the deep neural network in identifying complex patterns within the knee X-ray image dataset. Additionally, Batch normalization and Leaky RELU have been applied to all convolution layers. Batch normalization improves the recognition rate for test samples, and leaky [46,47] is a form of activation function that adds a small negative slope to the ReLU to maintain and propagate the weight updates throughout the entire propagation process [48,49]. Max pooling has been used in the pooling layer. The convolution layer result has been flattened further into three dense layers, followed by a dropout to remove unneeded features and provide for the softmax activation function. The Adamax optimization approach has been used to assemble each convolutional layer, and layers of dropout with different units have been employed, which implies that to

prevent overfitting on the training dataset, 25 % of neurons have to be arbitrarily assigned zero after every epoch. In the output layer, feature maps from different layers have been concatenated, and the result has been given to the output layer via a binary class (normal or osteoporosis).

Table 3. Details of custom CNN layers.

Layer type	Number of filters	Kernel size	Output shape	#trainable parameters
conv2d_61 (Conv2D)	32	(2x2)	127x255x32	416
batch_normalization_61 (BatchNormalization)	-	-	127x255x32	128
leaky_re_lu_88 (LeakyReLU)	-	-	127x255x32	0
max_pooling2d_60 (MaxPooling2D)	-	-	63x127x32	0
conv2d_62 (Conv2D)	64	(3x3)	61x125x64	18496
batch_normalization_62 (BatchNormalization)	-	-	61x125x64	256
leaky_re_lu_89 (LeakyReLU)	-	-	61x125x64	0
max_pooling2d_61 (MaxPooling2D)	-	-	20x41x64	0
conv2d_63 (Conv2D)	256	(3x3)	18x39x256	147712
batch_normalization 63 (BatchNormalization)	-	-	18x39x256	1024
leaky_re_lu_90 (LeakyReLU)	-	-	18x39x256	0
max_pooling2d_62 (MaxPooling2D)	-	-	6x13x256	0
conv2d_64 (Conv2D)	512	(2x2)	5x12x512	524800
batch_normalization 64 (BatchNormalization)	-	-	5x12x512	2048
leaky_re_lu_91 (LeakyReLU)	-	-	5x12x512	0
max_pooling2d_63 (MaxPooling2D)	-	-	2x6x512	0
flatten_13 (Flatten)	-	-	6144	0
dense_41 (Dense)	-	-	256	1573120
leaky_re_lu_92 (LeakyReLU)	-	-	256	0
dropout_10 (Dropout)	-	-	256	0
dense_42 (Dense)	-	-	128	32896
leaky_re_lu_93 (LeakyReLU)	-	-	128	0
dense_43 (Dense)	-	-	2	258

3. Experiment Evaluation and Results

The section highlights the ensemble model's predicted outcomes. Additionally, highlights the environment in which implementation has been carried out, model descriptions, and the results of all employed techniques.

3.1. *Implementation environment*

The Jupyter Notebook provided by Google's Collaboratory platform has been utilized for the entire experimental work. The platform provides the GPU-based 'free cloud service' for research purposes or scientific computations. A system combination with 13 GB RAM, 78 GB Hard Drive, 15 GB GPU, and NVIDIA T4 GPU has been utilised during the experimental procedure.

3.2. *Performance evaluation*

The diagnostic ability of the trained models has been assessed using a test dataset of knee x-ray images from both the normal and osteoporosis groups. To comprehensively evaluate the model's performance, accuracy and specificity have been used as measurement indicators. These parameters have been specified by the actual label and predicted label that comprise the confusion matrix, which are shown in Fig. 5. The model's performance has been clarified by a confusion matrix consisting of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Among them, S1 and S2 are two class labels; TP signifies that it has been proven to be an osteoporosis class, which is an osteoporosis class. FP denotes that it has been determined to be a normal class, but it is also a normal. TN implies that it is classified as an osteoporosis class. However, it is actually a normal class FN depicts that it is considered normal, yet it is actually an osteoporosis class. The confusion matrix for ResNet101, DenseNet169, VGG16, EfficientNetB5, custom CNN, and Ensemble are shown in Fig. 6.

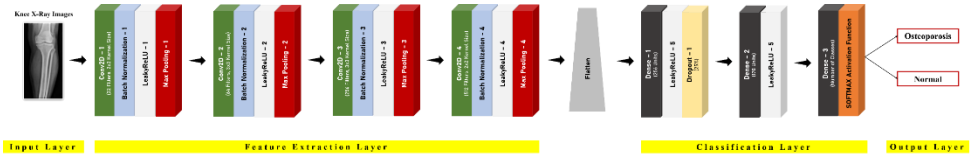

Fig. 4. Custom CNN architecture.

		Predicted	
	Binary class confusion matrix	S2	S2
Actual	S1	TP	FN
	S2	FP	TN

Fig. 5. Confusion matrix for Binary Class.

3.3. *Model description*

To avoid overfitting, a total of 100 epochs have been done when training the model with a batch size of 25 and utilizing early stopping. In task-specific binary classification, to differentiate osteoporosis from normal images in the test dataset, the binary_cross-entropy

loss function to compute cross-entropy losses between the predicted and target outcomes has been chosen. An optimizer is required to assemble a deep learning model. Optimizers are algorithms that execute all possible solutions iteratively until they reach a point that is optimum or satisfactory by altering attributes such as weight and learning rate to lower the overall loss and enhance accuracy [50]. It benefits in achieving quicker outcomes because a deep learning model often has millions of parameters. The proposed model has been compiled with the Adamax optimizer.

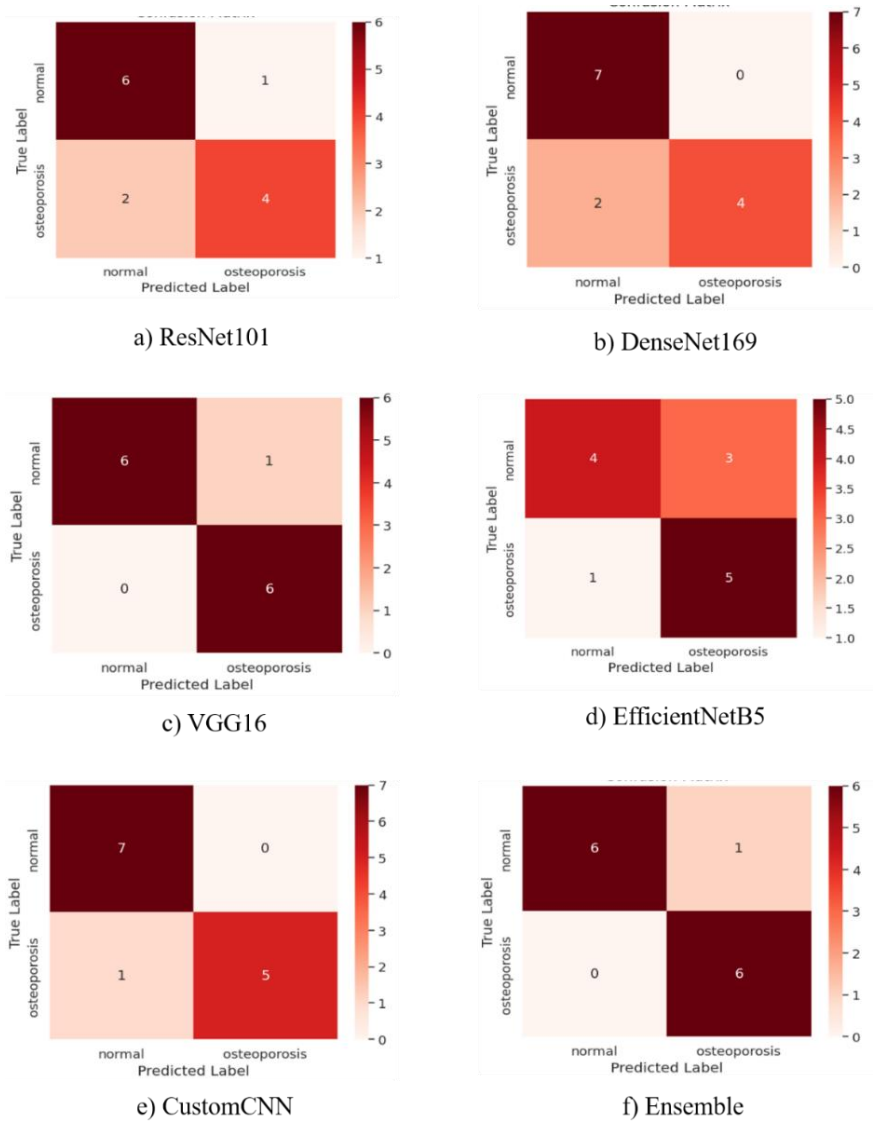

Fig. 6. Confusion matrix for a) ResNet101, b) DenseNet169, c) VGG16, d) EfficientNetB5, e) custom CNN, f) Ensemble.

AdaMax is an acronym that represents "adaptive moment estimation with maximum." It is an extended form of the gradient descent Adam optimizer that employs the maximum value of the second momentum portion and generalizes the approach to the infinite norm (max) [51]**.** The learning rate is a tuning parameter that determines the pace at which deep or machine learning models learn, as well as the number of steps needed to reduce the loss function's value [52]. The momentum parameter is utilized to enhance both the speed and accuracy of model training. A distinct function known as the activation function determines a node's output.

3.4. *Experimental Results*

The section discusses the ensemble methodology's outcomes from testing. Using knee X-ray images, experiments have been conducted to predict and categorize confirmed osteoporosis cases. The model's efficacy has been assessed using performance parameters, and the model has been evaluated for binary class (osteoporosis and normal) data. Table 4 reports the aggregate accuracy of the customized CNN model, pretrained models and ensemble model and also deduces the model's effectiveness. Loss curves for pretrained models, custom CNN model, and ensemble model for testing data are shown in Fig. 7. As shown in Fig. 7, there is a positive trend toward improving accuracy and reducing loss.

Table 4. Performance of all applied models.

Models	Accuracy (%)	Specificity (%)
ResNet101	76.92	66.67
DenseNet169	84.61	66.67
Vgg16	91.83	83.33
EfficientNetB5	69.23	57.14
Custom CNN	92.30	85.71
Ensemble Approach	95.8	100

3.5. *Comparative analysis*

Numerous works from the literature have been reviewed, including the current state-of-the-art for osteoporosis detection and classification. A comparison of existing techniques for predicting osteoporosis is shown in Table 5, and the proposed system achieved the highest accuracy of 100 %. It is evident from the table that the ensemble methodology has outperformed all current models for osteoporosis detection and classification in terms of accuracy. The ensemble methodology outlined in this paper has functioned more precisely to aid rheumatologists in the more efficient detection of osteoporosis.

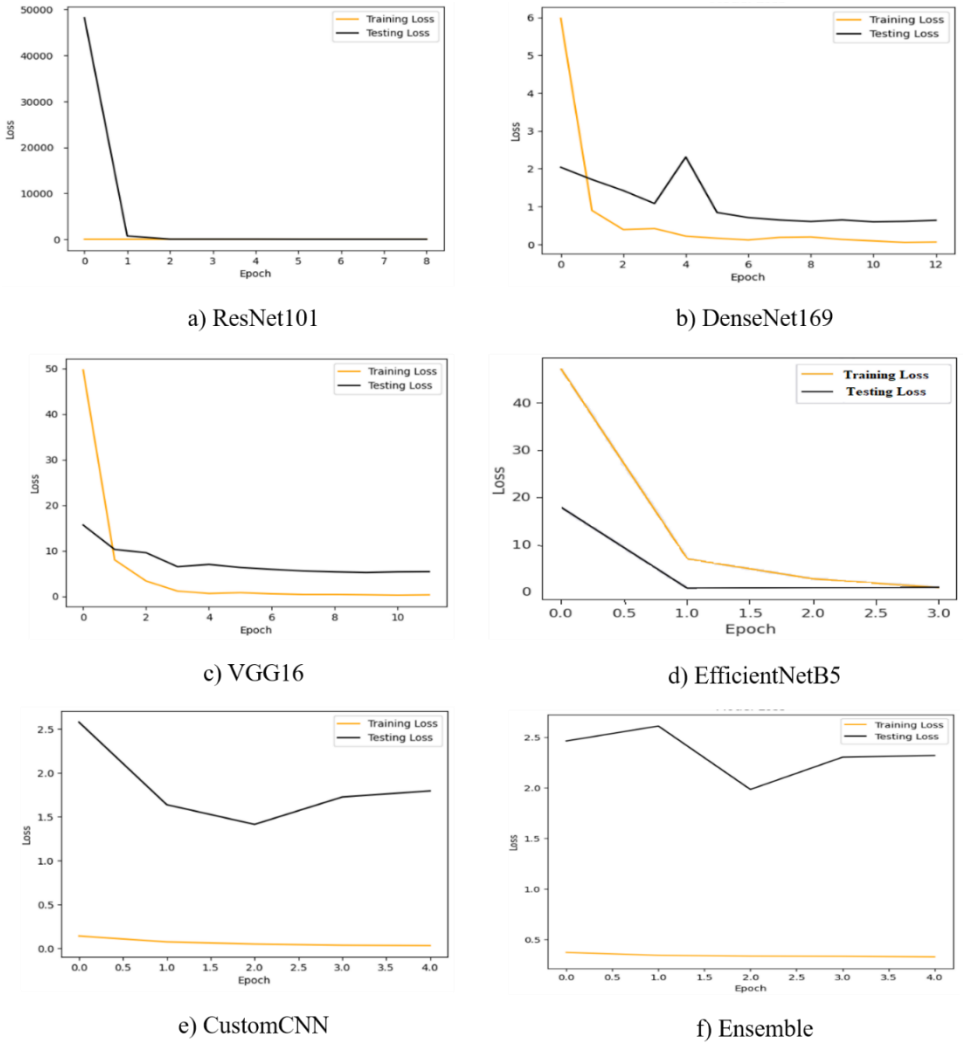

Fig. 7. Loss curve for a) ResNet101, b) DenseNet169, c) VGG16, d) EfficientNetB5, e) custom CNN, f) Ensemble.

Table 5. Comparative analysis.

Authors & Years	Techniques Used	Imaging Modality	Performance
Tejaswini *et al.* 2016 [53]	Feed-forward backpropagation	CT	-
Tecle *et al.* 2019 [54]	CNN[VGG16-TR-TF]	X-RAY	Accuracy- 84 %
Jang *et al.* 2021 [34]	VGG16 and Grad-CAM	X-RAY	Accuracy- 71.8 % and AUC- 70 %.
Tang *et al.* 2019 [55]	CNN	CT images	Accuracy- 76.5 %

Meng *et al.* 2019 [56]	ANN and OSTA	X-RAY and dual-energy X-ray absorptiometry	Accuracy- 78.8 %
Ferizi *et al.* 2019 [57]	ANN	MRI	Accuracy- 71 %
Nam *et al.* 2019 [58]	GLCM	CT	Accuracy- 92.5 %, precision- 93.9 % recall- 0.969, F1 score- 95.4, AUC- 90 %
Yamamoto *et al.* 2020 [59]	ResNet18, ResNet34, GoogleNet, EfficientNetB3, and EfficientNetB4	Hip Radiographs	Accuracy- 79 %, Sensitivity- 86 %, Specificity-86 %
Mao *et al.* 2022 [60]	Convolutional Neural Network	Lumbar Radiographs	AUC- 78.5 %
Sato *et al.* 2022 [61]	Deep Learning	X-rays	AUC- 89 %
Abubakar *et al.* 2022 [62]	VGG16	X-rays	Accuracy- 88 %, sensitivity- 90 %, specificity-91 %
Dodamani & Danti, 2023 [63]	VGG16, VGG19, DenseNet121, Resnet50, and InceptionV3	X-rays	Accuracy- 93.4 %
Mane *et al.* 2023 [64]	VGG16, InceptionV3, ResNet50	X-rays	Accuracy- 95 %
Naguib *et al.* 2024 [65]	AlexNet and ResNet50	X-rays	Accuracy- 85.42 %
Our Work	**Ensemble of ResNet101, DenseNet169, VGG16 EfficientNetB5 and custom CNN**	**X-rays**	**Accuracy- 95.8 %**

4. Discussion

Osteoporosis is an undiagnosed bone disorder that affects millions of individuals and harms healthcare systems, society, and the affected individuals worldwide. A physician utilizes X-ray diagnostic procedures to treat bone problems since they are the most widely used method of detecting bone mineral density. These methods are inadequate for screening osteoporosis due to their high price, high radiation dosage, huge size, and limited availability of specialized infrastructure. Due to these drawbacks, there has been increasing interest in developing computer-aided or automated screening tools for the detection of osteoporosis. Accurate and timely diagnosis of osteoporosis would lead to improved clinical care, including prevention and appropriate pharmacological or surgical treatment.

With the rise of digital information, artificial intelligence technologies have regained popularity and demonstrated potential in many areas of bone disorders. Deep learning algorithms can relieve radiologists and pathologists of tedious and repetitive work by ingesting, analyzing, and reporting massive amounts of data across multiple modalities to

diagnose illnesses and guide therapeutic choices. Among deep learning methods, CNN has performed more rapidly and with higher classification accuracy. the researchers utilized the advantage of CNN networks' capacity to categorize knee X-ray images by analyzing the normal and osteoporosis class images and then automatically classifying them. The goal of the study is to build an ensemble model that can predict osteoporosis based on patient X-ray scans. Early osteoporosis case prediction is possible using the proposed clinical decision-making framework. The osteoporosis identification utilizing knee X-ray images saves time for tool training on a dataset gathered from numerous databases and radiologists may find value in these trained tools. To evaluate the trained models' diagnostic ability, x-ray images were utilized as the test set. The evaluation parameters, such as specificity and accuracy, are calculated to provide an in-depth assessment of the test dataset screening performance.

To demonstrate the accuracy and significance of research work, the ensemble model (shown in Table 5) has been compared to other recently developed models and the substantial literature on osteoporosis identification. The previously proposed deep learning models for the detection of osteoporosis had achieved good performance but still, but there is still a need for improvement. According to the findings of the Ensemble methodology, deep learning with CNNs may have significant effects on the automatic detection and classification of knee X-ray images related to osteoporosis diagnosis. Integration of deep learning technologies in osteoporosis care could ameliorate the accuracy and speed of diagnosis, assist clinical decision-making, and lead to better health outcomes.

5. Conclusion

Accurate and timely diagnosis of osteoporosis is crucial to prevent widespread osteoporotic fractures. In this paper, the researchers have developed an ensemble method for predicting osteoporosis. A database of knee X-ray images from individuals with verified cases and normal cases is used to classify osteoporosis automatically. There are 360 images in the dataset, 180 of which are normal cases and 180 are osteoporosis proven. A total accuracy of 95.8 % is achieved, which is better than others related to osteoporosis detection. The suggested technique outperforms existing deep learning methods in terms of specificity and accuracy.

The ensemble model outperforms previous cutting-edge techniques in osteoporosis diagnosis. Also, the ensemble technique is computationally efficient because of a few parameters. After evaluating the proposed network's performance score, it may be utilized as a benchmark in osteoporosis screening. The ensemble model might be utilized as a supplement tool for doctors to diagnose osteoporosis using knee X-ray pictures.

The future research will concentrate on working with larger datasets or for osteoporosis screening, enhancing the method's robustness. Current research might be further enhanced by developing automated methods for osteoporosis prediction based on multidisciplinary deep learning models and images, such as digital CT scans, MRIs, etc. The work also

encourages the researchers to focus on different validation methods and cross-datasets for network training and testing.

References

1. M. Kastner, L. Perrier, S. E. P. Munce, C. C. Adhihetty, and A. Lau, Osteoporos Int. **29**, 5 (2017). https://doi.org/10.1007/s00198-017-4248-0
2. C. M. Court-brown and B. Caesar, Int. J. Care Inj. **37**, 691 (2006). https://doi.org/10.1016/j.injury.2006.04.130
3. R. Mallina, N. K. Kanakaris, and P. V Giannoudis, Knee **17**, 181 (2010). https://doi.org/10.1016/j.knee.2009.10.011
4. F. Pouresmaeili, B. Kamalidehghan, M. Kamarehei, and Y. M. Goh, Dove Press **2018**, 2029 (2018). https://dx.doi.org/10.2147/TCRM.S138000
5. S. Castañeda, A. Casas, A. González, G. Martínez, D. Guerra, and X. Nogués, Clin. Transl. Oncol. **24**, 2090 (2022). https://doi.org/10.1007/s12094-022-02872-1
6. K. A. Patil, K. V. M. Prashanth, and A. Ramalingaiah, Int. J. Res. Orthop. **7**, 4 (2021). https://dx.doi.org/10.18203/issn.2455-4510
7. A. Inui, H. Nishimoto, Y. Mifune, T. Yoshikawa, and I. Shinohara, Bioengineering **10**, ID 277 (2023). https://doi.org/10.3390/bioengineering10030277
8. J. Smets, E. Shevroja, T. Hügle, W. D. Leslie, and D. Hans, J. Bone Miner. Res. **36**, 833 (2021). https://doi.org/10.1002/jbmr.4292
9. D. Shao, I. Dai, N. Li, X. Cao, W. Zhao, L. Cheng, Z. Rong, L. Huang, Y. Wang, and J. Zhao, Brief Bioinform. **23**, ID bbab523 (2022). https://doi.org/10.1093/bib/bbab523
10. R. Yang and Y. Yu, Front. Oncol. 11, ID 638182 (2021). https://doi.org/10.3389/fonc.2021.638182
11. I. D. Apostolopoulos and T. A. Mpesiana, Phys. Eng. Sci. Med. **43**, 635 (2020). https://doi.org/10.1007/s13246-020-00865-4
12. S. S. Yadav and S. M. Jadhav, J. Big Data **6**, ID 113 (2019). https://doi.org/10.1007/s13246-020-00865-4
13. D. R. S. R. V Kulkarni, Evol. Intell. **15**, 1 (2022). https://doi.org/10.1007/s12065-020-00540-3
14. S. Iqbal, A. N. Qureshi, J. Li, and T. Mahmood, Arch Computat. Methods Eng. **30**, 3173 (2023). https://doi.org/10.1007/s11831-023-09899-9
15. F. Emmert-streib, Z. Yang, H. Feng, S. Tripathi, and M. Dehmer, Front. Artif. Intell. **3**, ID 4 (2020). https://doi.org/10.3389/frai.2020.00004
16. B. Bhinder, C. Gilvary, N. S. Madhukar, and O. Elemento, Cancer Discov. **11**, 900 (2021). https://doi.org/10.1158/2159-8290.CD-21-0090
17. Z. Ardalan and V. Subbian, Front. Artif. Intell. **5**, ID 780405 (2022). https://doi.org/10.3389/frai.2022.78040
18. A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, J. Big Data **9**, ID 102 (2022). https://doi.org/10.1186/s40537-022-00652-w
19. I. M. Wani and S. Arora, Proceedings of ICRIC 2019, Lect. Notes Electr. Eng. **597**, 65 (2019). https://doi.org/10.1007/978-3-030-29407-6_6
20. T. K. Yoo, S. K. Kim, E. Oh, and D. W. Kim, Adv. Comput. Intell. IWANN 2013, Lect. Notes Comput. Sci. **7903**, 181 (2012). https://doi.org/10.1007/978-3-642-38682-4_21
21. W. O. Yang, C. Lai, M. Tsou and L. Hwang, Int. J. Environ. Res. Public Health **18**, 7635 (2021). https://doi.org/10.3390/ijerph18147635
22. S. Dadsetan, G. Kitamura, D. Arefan, Y. Guo, L. Yang, and S. Wu, medRxiv (2022). https://doi.org/10.1101/2022.02.03.22270400
23. T. Iliou, C. Anagnostopoulos, and G. Anastassopoulos, Int. J. Artif. Intell. **23**, ID 14500146 (2014). https://doi.org/10.1142/S0218213014500146
24. J. Scanlan, F. F. Li, O. Umnova, G. Rakoczy, N. Lövey, and P. Scanlan, Bioengineering **5**, ID 107 (2018). https://doi.org/10.3390/bioengineering5040107

25. H. K. Lim, H. Il, H. Id, S. P. Id, and J. Han, PLoS ONE **16**, ID e0247330 (2021). https://doi.org/10.1371/journal.pone.0247330
26. L. Liu, M. Si, H. Ma, M. Cong, Q. Xu et al., BMC Bioinform. **23**, ID 63 (2022). https://doi.org/10.1186/s12859-022-04596-z
27. S. Bhattacharya, D. Nair, and A. Bhan, 6th Int. Conf. Signal Process. Integr. Networks (2019).
28. S. Sukegawa, A. Fujimura, A. Taguchi, N. Yamamoto, and A. Kitamura, Sci. Rep. **12**, ID 6088 (2022). https://doi.org/10.1038/s41598-022-10150-x
29. K. Lee, S. Jung, J. Ryu, S. Shin, and J. Choi, J. Clin. Med. **9**, ID 392 (2020). https://doi.org/10.3390/jcm9020392
30. M. Tassoker, M. U. Ozic, and F. Yuce, Res. Square (2022). https://doi.org/10.21203/rs.3.rs-1245777/v1
31. D. Devikanniga, Int. J. Intell. Networks **1**, 43 (2020). https://doi.org/10.1016/j.ijin.2020.05.004
32. S. T. Gokdeniz and K. Kamburoğlu, World J. Radiol. **14**, 55 (2022). https://dx.doi.org/10.4329/wjr.v14.i3.55
33. R. Gaudin, W. Otto, I. Ghanad, S. Kewenig, C. Rendenbach, V. Alevizakos, P. Grün, F. Kofler, M. Heiland, and C. V. See, Med. Sci. **12**, ID 49 (2024). https://doi.org/10.3390/medsci12030049
34. R. Jang, J. H. Choi, N. Kim, J. S. Chang, P. W. Yoon, and C. H. Kim, Sci. Rep. **11**, 19997 (2021). https://doi.org/10.1038/s41598-021-99549-6
35. S. Kumar and R. Rani, Int. J. Imaging Syst. Technol. **32**, 1464 (2022). https://doi.org/10.1002/ima.22770
36. U. Bashir, K. Singh and V. Mansotra, J. Sci. Res. **17**, 195 (2025). https://dx.doi.org/10.3329/jsr.v17i1.74640
37. Osteoporosis Knee X-ray Dataset | Kaggle. https://www.kaggle.com/datasets/stevepython/osteoporosis-knee-xray-dataset
38. A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, Artfic. Intell. Rev. **53**, 5455 (2020). https://doi.org/10.1007/s10462-020-09825-6
39. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan et al., J. Big Data **8**, ID 53 (2021). https://doi.org/10.1186/s40537-021-00444-8
40. M. M. Taye, Computation **11**, ID 52 (2023). https://doi.org/10.3390/computation11030052
41. I. H. Sarker, SN Comput. Sci. **2**, ID 420 (2021). https://doi.org/10.1007/s42979-021-00815-1
42. A. Ramdan, A. Heryana, and A. Arisal, IEEE Xplore (2020). https://doi.org/10.1109/ICRAMET51080.2020.9298575
43. A. Vulli, P. N. Srinivasu, M. S. K. Sashank, J. Shafi, J. Choi, and M. F. Ijaz, Sensors **22**, ID 2988 (2022). https://doi.org/10.3390/s22082988
44. G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, IEEE Xplore (2017). https://doi.org/10.1109/CVPR.2017.243
45. W. A. Hamwi and M. M. Almustafa, Informatics Med. Unlocked **32**, ID 101004 (2022). https://doi.org/10.1016/j.imu.2022.101004
46. X. Zhang, Y. Zou, and W. Shi - *22nd Int. Conf. on Digital Signal Processing (DSP)* (2017).
47. A. Mujhid, S. Surono, N. Irsalinda, and A. Thobirin, Fuzzy Systems and Data Mining (2022) pp. 50–57.
48. S. Ioffe and C. Szegedy, arXiv (2015). https://doi.org/10.48550/arXiv.1502.03167
49. T. Jiang and J. Cheng – *Int. Conf. on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, (Shijiazhuang, China, 2019).
50. F. Mehmood, S. Ahmad, and T. K. Whangbo, Mathematics **11**, ID 1360 (2020). https://doi.org/10.3390/math11061360
51. D. Yi, J. Ahn, and S. Ji, Appl. Sci. **10**, ID 1073 (2020). https://doi.org/10.3390/app10031073
52. K. Nakamura, B. Derbel, K. Won, and B. Hong, Electronics **10**, ID 2029 (2021). https://doi.org/10.3390/electronics10162029
53. E. Tejaswini, P. Vaishnavi, and R. Sunitha, IEEE Xplore (2016). https://doi.org/10.1109/ICACCI.2016.7732272
54. N. Tecle, J. Teitel, M. R. Morris, N. Sani, D. Mitten, and W. C. Hammert, J. Hand Surg. Am. **45**, 175 (2019). https://doi.org/10.1016/j.jhsa.2019.11.019

55. C. Tang, W. Zhang, H. Li, L. Li, Z. Li, A. Cai, L. Wang, D. Shi and B. Yan, arXiv (2019). https://doi.org/10.48550/arXiv.1910.06777
56. J. Meng, N. Sun, Y. Chen, Z. Li, X. Cui et al., J. Int. Med. Res. **47**, 3088 (2019). https://doi.org/10.1177/0300060519850648
57. U. Ferizi, H. Stephen, and C. Gregory, J. Magn. Reson. Imaging **31**, 4, 368 (2020). https://doi.org/10.1097/BOR.0000000000000607
58. K. H. Nam, I. Seo, D. H. Kim, J. Il Lee, B. K. Choi, and I. H. Han, J. Kor. Neurosurg. Soc. **62**, 442 (2019). https://doi.org/10.3340/jkns.2018.0178
59. N. Yamamoto, S. Sukegawa, A. Kitamura, R. Goto et al., Biomolecules **10**, ID 1534 (2020). https://doi.org/10.3390/biom10111534
60. L. Mao, Z. Xia, L. Pan, J. Chen, X. Liu, Z. Li, Z. Yan, G. Lin, H. Wen, and B. Liu, Front. Endocrinol. **13**, ID 971877 (2022). https://doi.org/10.3389/fendo.2022.971877
61. Y. Sato, N. Yamamoto, N. Inagaki, Y. Iesaki, T. Asamoto, T. Suzuki, and S. Takahara, Biomedicines **10**, ID 2323 (2022). https://doi.org/10.3390/biomedicines10092323
62. U. B. Abubakar, M. M. Boukar, and S. Adeshina, Int. J. Adv. Comput. Sci. Appl. **13**, 246 (2022). https://doi.org/10.14569/IJACSA.2022.0130829
63. P. S. Dodamani and A. Danti, Int. J. Online Biomed. Eng. **19**, 66 (2023). https://doi.org/10.3991/ijoe.v19i08.39235
64. P. R. Mane, J. Vemulapalli, N. S. Reddy, N. Anudeep, and G. Prabhu, J. Phys. Conf. Ser. **2571**, ID 012017 (2023). https://doi.org/10.1088/1742-6596/2571/1/012017
65. S. M. Naguib, M. K. Saleh, H. M. Hamza, K. M. Hosny, and M. A. Kassem, Sci. Rep. **14**, ID 25434 (2024). https://doi.org/10.1038/s41598-024-75549-0