

Article

Hybrid SFNet Model for Bone Fracture Detection and Classification Using ML/DL

Dhirendra Prasad Yadav ¹, Ashish Sharma ¹, Senthil Athithan ² , Abhishek Bhola ² , Bhisham Sharma ^{3,*} 
and Imed Ben Dhaou ^{4,5,6,*} 

¹ Department of Computer Engineering and Applications, GLA University, Mathura 281406, Uttar Pradesh, India

² Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, Andhra Pradesh, India

³ Chitkara University School of Engineering and Technology, Chitkara University, Himachal Pradesh 174103, India

⁴ Department of Computer Science, Hekma School of Engineering, Computing and Informatics, Dar Al-Hekma University, Jeddah 22246-4872, Saudi Arabia

⁵ Department of Computing, University of Turku, 20500 Turku, Finland

⁶ Higher Institute of Computer Sciences and Mathematics, Department of Technology, University of Monastir, Monastir 5000, Tunisia

* Correspondence: bhisham.pec@gmail.com (B.S.); imed.bendhaou@utu.fi (I.B.D.)

Abstract: An expert performs bone fracture diagnosis using an X-ray image manually, which is a time-consuming process. The development of machine learning (ML), as well as deep learning (DL), has set a new path in medical image diagnosis. In this study, we proposed a novel multi-scale feature fusion of a convolution neural network (CNN) and an improved canny edge algorithm that segregate fracture and healthy bone image. The hybrid scale fracture network (SFNet) is a novel two-scale sequential DL model. This model is highly efficient for bone fracture diagnosis and takes less computation time compared to other state-of-the-art deep CNN models. The innovation behind this research is that it works with an improved canny edge algorithm to obtain edges in the images that localize the fracture region. After that, grey images and their corresponding canny edge images are fed to the proposed hybrid SFNet for training and evaluation. Furthermore, the performance is also compared with the state-of-the-art deep CNN models on a bone image dataset. Our results showed that SFNet with canny (SFNet + canny) achieved the highest accuracy, F1-score and recall of 99.12%, 99% and 100%, respectively, for bone fracture diagnosis. It showed that using a canny edge algorithm improves the performance of CNN.

Keywords: bone fracture; hybrid; CNN; canny; fusion; classification; X-ray



Citation: Yadav, D.P.; Sharma, A.; Athithan, S.; Bhola, A.; Sharma, B.; Dhaou, I.B. Hybrid SFNet Model for Bone Fracture Detection and Classification Using ML/DL. *Sensors* **2022**, *22*, 5823. <https://doi.org/10.3390/s22155823>

Academic Editor: Jose Manuel Molina López

Received: 10 July 2022

Accepted: 2 August 2022

Published: 4 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bone is an essential component of the human body. It helps people to move from one place to another place. A bone fracture may happen due to an accident or other reason. After a fracture, treatment is necessary as early as possible. An orthopedic surgeon examines an X-ray or CT-scan image to detect fractured bone. After that, treatment is given to the patient. The treatment success depends on an orthopedic surgeon's experience and expertise. This manual process is time consuming and the availability of experts in a remote area is less. In recent days, diagnosis of real-time medical problems is done by machine learning and deep learning technique [1,2]. In addition, several deep CNN models were reported in the past to demonstrate its effectiveness for various applications [3,4].

In the current research work, we aim to detect the bone fracture with a high level of accuracy so that the proper identification of the fracture can be performed and the severity of the bone fracture can be judged. The X-ray images are more sensitive to noise that can be removed using the image processing technique. Several pieces of the literature are

available for the analysis of different types of human bones. In this regard, Lu et al. [5] designed a universal fracture detection system using the deep CNN techniques. First, the quality of an image is improved using preprocessing techniques. After that, data augmentation is used to enlarge the size of the dataset. Finally, Ada-ResNeSt is used for the classification of fractured and healthy bone with a mean precision of 68.4%. In other research, Guan et al. [6] enhanced their model performance on the arm bone X-ray images. They designed a model for fracture detection in arm bones. Three factors are included in the core changes. First, to gain more fractural details, a new back-bone network is developed based on the function pyramid architecture. Second, to improve the contrast of original images, an image preprocessing method involves an opening procedure and pixel value transformation.

All the methods discussed above and in Section 2 have done remarkable work in classifying the fractured bone. However, most research is based on machine learning techniques requiring handcrafted training features. This manual feature extraction method requires expertise and classification performance is not optimal. Several pieces of research have been reported based on the transfer learning approach, but their performance is not optimal. In this study, we have trained and analyzed the performance of the AlexNet, VGG16, ResNeXt and MobileNetV2 on the X-ray image of the bone image dataset. The performance of these models is remarkable for the healthy bone, but lacks in the fracture bone diagnosis. The major contribution of this research is we have designed a new multimodal feature fusion-based hybrid SFNet model that takes less computation time, and efficiency is high compared to other state-of-the-art CNN models. The grey images and their corresponding canny image are fed to the model for training and validation. Furthermore, hybrid SFNet + canny achieved the highest classification accuracy. It showed that using a canny edge algorithm improves the performance of CNN.

1.1. Problem Statement

Bone is an essential component of the human body. It helps people to move from one place to another place. A bone fracture may happen due to an accident or other reason. After a fracture, treatment is necessary as early as possible. An orthopedic surgeon examines an X-ray or CT-scan image to detect fractured bone. The diagnosis process for fractured bone requires expertise. Hence, a robust and efficient automated system needs to be designed that can diagnose bone fractures in real-time to assist an orthopedic surgeon. In the past, several research studies were conducted to design an automated system using ML and DL. The ML technique required hand-crafted features due to this optimal solution not being available. In contrast, the DL method extracts features automatically and provides an enhanced diagnosis, but its computation cost is high and requires a large dataset to train the model.

1.2. Motivation

In recent days, the biomedical domain has been tuned with the machine learning and deep learning models. Bone diagnosis using artificial intelligence techniques has been used in past research. These methods have done a tremendous job, but lack in efficiency, and computation cost is high. In the current work, a novel hybrid SFNet has been developed which is highly efficient and the computation cost is less. The hybrid SFNet is a two-scale DL model, which has less neurons, due to this computation cost being less and the high efficiency that is achieved by the training of the model at two scales. First, localization of fractured bone is performed using an improved canny edge detection technique. After that, grey image and their corresponding canny image are fed into hybrid SFNet for deep feature extraction. This process significantly improves the bone fracture diagnosis.

The major contribution of the proposed study is as follows.

- (1) Designed a novel hybrid feature fusion-based SFNet for bone diagnosis.
- (2) The canny edge image improves the performance.
- (3) The experiment is conducted on a publicly available dataset.
- (4) The proposed model classification performance for fractured bone is highest.

1.3. Organization of the Research

Section 2 describes the overview of the past research on bone diagnosis. In Section 3, the architecture of different deep CNN models has been discussed, and in Section 4, the proposed model architecture is explained. The comparisons of results are described in Section 5. In Section 6, a summary of the current research work is discussed. Finally, the proposed method is concluded in Section 7.

2. Related Work

Basic machine learning techniques, such as preprocessing and feature extraction, has been used in this study [7]. The relaxed digital straight line (RDSS) approach for restoring the false contour to overcome the error that occurs at the time of segmentation in the contour is elaborated in [8]. The same author, Basha et al. [9] developed a model based on the machine learning technique for bone fracture diagnosis. Firstly, the Gaussian filter is used to improve the quality of an X-ray image. Then, the edge is detected using the canny edge algorithm. Finally, the Harris corner detection method is used to identify the fractured regions. This method identifies fractured bone with an accuracy of 92%. Feature extraction is improved with a variety of image processing techniques, such as controllable filters, image projection integration, and image pixel intensity [10]. All of these methods have done a tremendous job for bone diagnosis. However, the performance of these methods is not optimal due to the handcrafted features used for the training of the model.

Recently, several researchers reported that merging the deep CNN model to form a pool of features improves the overall classification performance. Recently, Lin et al. [11] developed fracture R-CNN (Region-based Convolutional Neural Network) for skull fracture detection. They employed prior clinical knowledge in faster R-CNN to enhance the classification performance. Kitamura et al. [12] developed an ensemble-based deep CNN model for ankle fracture detection. Their method includes Xception, InceptionV3 and ResNet for feature extraction. The ensemble-based model classifies healthy and fractures with an accuracy of 81%. In the transfer learning-based approach, Choi et al. [13] applied YOLOV3 model for skull fracture identification and achieved a sensitivity of 91.7%. In similar research, Kitamura et al. [14] and Kim et al. [15] applied pre-trained DenseNet-121 and InceptionV3 models for bone diagnosis. The method described in [14] achieved 95% accuracy, whereas the method described in [15] achieved an accuracy of 95.4%.

Some research reported that they first trained their model on a bone image dataset before classification is performed. In this regard, Yang et al. [16] developed two deep CNN models for intertrochanteric fracture segmentation and identification. They split the dataset into two parts, i.e., training and testing, in which 32,045 and 11,465 images are used, respectively. First, the region of interest (ROI) is extracted using cascade architecture-based CNN. After that, another CNN is used for segmentation and identification. In other research, Haitaamar et al. [17] performed preprocessing on a rib fracture image to resize the image into $128 \times 128 \times 333$ pixels. After that, a semantic segmentation technique was applied to locate the fractured region of the ribs. Finally, the UNet model was used to classify CT scan images and achieved an accuracy of 88.54%. Nguyen et al. [18] applied the Yolo 4 model for the localization of fractured bone. The data augmentation technique was also applied and their model performance was compared with original and augmented datasets. On an original dataset, this method can differentiate fracture and healthy regions with an accuracy of 81.94%. A pyramid network was designed by Wang et al. [19] for bone diagnosis. They trained the pyramid network on an X-ray image. Their method achieved an accuracy of 88.7%. Ma et al. [20] identified bone fracture in two steps; in the first step faster R-CNN was utilized to identify 20 fracture regions and in the second step, a novel CrackNet for bone fractured classification. Their method classifies healthy and fractured bone with a classification accuracy of 90.14%. In similar research, a new Parallel Net approach for classifying fractured bone has been proposed using the two-scale method proposed by Wang et al. [21].

Yahalomi et al. [22] and Abbas et al. [23] applied pre-trained faster R-CNN on a small dataset and achieved an accuracy of 96% and 97%, respectively. Luo et al. [24] applied their expert knowledge and designed a decision tree for fractured bone diagnosis. Their method achieved an accuracy of 86.57%. Beyaz et al. [25] and Jones et al. [26] applied deep CNN for the feature extraction from bone images. Their method validation accuracy is 83% and 97.4%, respectively. Dupuis et al. [27] designed the Rayvolve[®] model for bone fracture detection in children. The external validation method classifies healthy and fractured bone with an accuracy of 95%. Hardalaç et al. [28] applied an ensemble-based deep CNN model for wrist fracture identification. Their method segregates fracture and healthy wrist bone with an accuracy of 86.39%.

Pranata et al. [29] classified calcaneus fractures in CT images using a fusion of the deep features and SURF features. They compared the performance of VGG16 and ResNet, which are two pre-trained deep CNN models. Out of these two models, ResNet achieved the highest classification accuracy of 98%. Mutasa et al. [30] applied the generative adversarial network (GAN) and digitally reconstructed radiographs (DRRs) techniques. The image dataset is created using GAN and classification is performed using a deep CNN model. Weikert et al. [31] applied a deep learning-based CNN model and achieved a classification accuracy of 90.2% for rib fracture identification. Tanzi et al. [32] applied the InceptionV3 model for the classification of proximal femur X-ray images. Their method classifies healthy and fractured femur images with an accuracy of 86%. In similar research, Lotfy et al. [33] applied DenseNet for femur fracture classification and achieved an accuracy of 89%. In [34], the authors have compiled a review work and defined the most important points, what should be taken into account to achieve that goal, and they have compared various models to humans in the classification of fractures. Table 1 provides a summary of the recent research for bone diagnosis.

Table 1. Summary of recent research for bone diagnosis.

Study	Type of Image	Feature Extraction Model	Dataset	Accuracy
Kitamura et al. [14]	X-ray	DenseNet-121	14,374 images	95%
Kim et al. [15]	X-ray	Pretrained InceptionV3	11,112 images	95.4%
Yang et al. [16]	CT Scan	CNN	43,510 images	89.4%
Haitamar et al. [17]	CT Scan	UNet	150 images	88.54%
Nguyen et al. [18]	X-ray	YoLo4	4405 images	81.91%
Wang et al. [19]	X-ray	Pyramid Network	3842 images	88.7%
Ma et al. [20]	X-ray	CrackNet	1052 images	90.14%
Wang et al. [21]	X-ray	ParallelNet	3842 images	87.8%
Yahalomi et al. [22]	X-ray	Faster R-CNN	38 images	96%
Abbas et al. [23]	X-ray	Faster R-CNN	50 images	97%
Luo et al. [24]	X-ray	Medical decision trees	1000 images	86.57%
Beyaz et al. [25]	X-ray	Deep CNN	2106 images	83%
Jones et al. [26]	X-ray	Deep CNN	715,343 images	97.4%
Dupuis et al. [27]	X-ray	Rayvolve [®]	5865 images	95%
Hardalaç et al. [28]	X-ray	Ensembles deep CNN	569 images	86.39%
Pranata et al. [29]	CT Scan	ResNet + VGG16 + SURF	1931 images	98%
Mutasa et al. [30]	X-ray	GAN + DRS	9063 images	96%
Weikert et al. [31]	CT Scan	Deep CNN	511 images	90.2%
Tanzi et al. [32]	X-ray	Inception V3	2453 images	86%
Lotfy et al. [33]	X-ray	DenseNet	1347 images	89%

3. Existing Methods

In the proposed research, we have analyzed the performance of AlexNet, MobileNetV2, ResNeXt and VGG16.

3.1. AlexNet Model

The AlexNet [35] model contains three dense layers and five convolutional layers. The first step is to scale the bone image to 256×256 pixels, corresponding to width and height, and a three-color channel that represents the depth of the image. The output of neurons is calculated using their respective weights as the inner product of a small portion of the image. First convolution layer (Conv1) size is fixed to 96 and a filter of size 11×11 with steps of four is added. After that, the output of the first convolutional layer is fed to the second convolution layer (Conv2) of 256 and filter of size 5×5 . The third (Conv3) and fourth layers (Conv4) have 384 kernels with filters of 3×3 . Finally, the last convolutional layers have a kernel of 256 and a filter of size 3×3 . A m-pooling of 2×2 is applied after each convolution layer. The FC1 and FC2 are the fully connected layer search with 4096 neurons. The FC3 uses the Softmax function to classify bone into healthy and fracture. The detailed architecture of the AlexNet for bone diagnosis is shown in Figure 1.

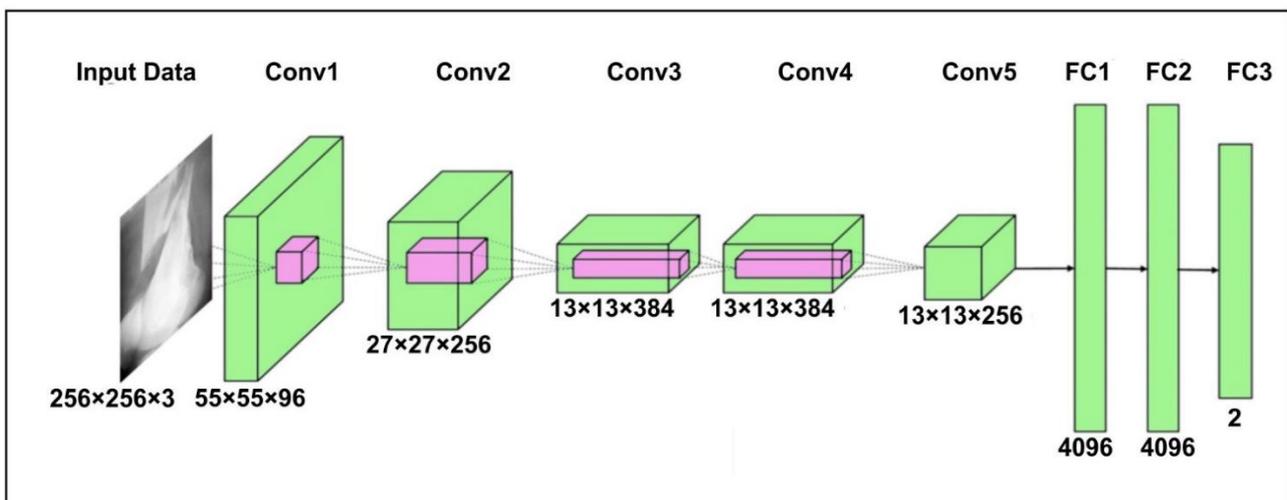


Figure 1. The AlexNet model architecture for bone diagnosis.

3.2. MobileNetV2 Model

In the proposed approach, we have used pre-trained MobileNetV2 [36] architecture trained on ImageNet. This model is a lightweight network and the significant aim is to minimize the parameters so that the computation is done quickly. To minimize the parameters, it uses a separate layer of convolution. In addition, MobileNetV2 uses depth-wise separable convolutions that are supportable in a very effective manner for a smartphone [37]. A depth-separable convolution layer is the combination of a point convolution layer and a deep convolution layer. At each input path, the deep convolution layer performs a single convolution. After that, the point convolution layer compares the output of the deep convolution layer. It works on a new communication mode that would link each of the current network layers to the previous network layers; with fewer convolution kernels, it can use the previous features repeatedly to create more feature maps. The basic structure has a growth rate of four with four tightly connected layers. Each layer considers input from its previous layer's output in the form of a feature map in this system. The architecture of the MobileNetV2 is shown in Figure 2.

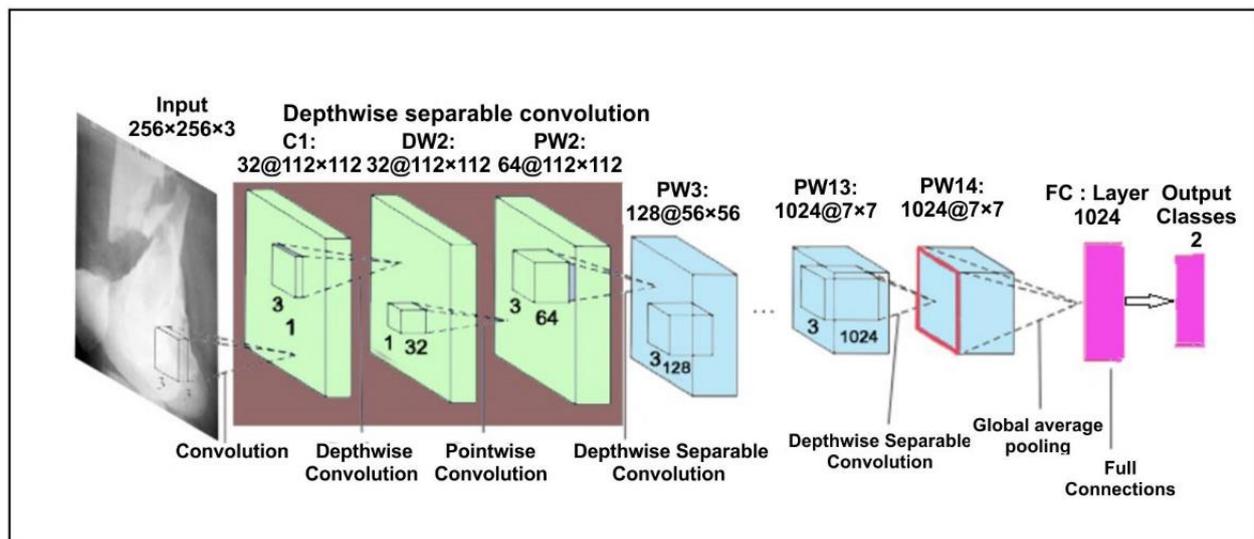


Figure 2. MobileNetV2 architecture for bone diagnosis.

3.3. ResNeXt Model

The ResNeXt is a residual class deep CNN model. This model is 50 layers deep and was first runner up in ILSVR 2016 image classification [38]. ResNeXt's major advantage is a reduction in the number of hyperparameters compared to ResNet-50 using cardinality by adding additional depth and width. The complete architecture is shown in Figure 3. Residual blocks in the model are defined as

$$\alpha = \beta + \sum_{k=1}^C F_k(\beta) \quad (1)$$

where, $C = 32$ (cardinality hyper-parameter), $\beta =$ input parameter vector, $\alpha =$ Output of the model.

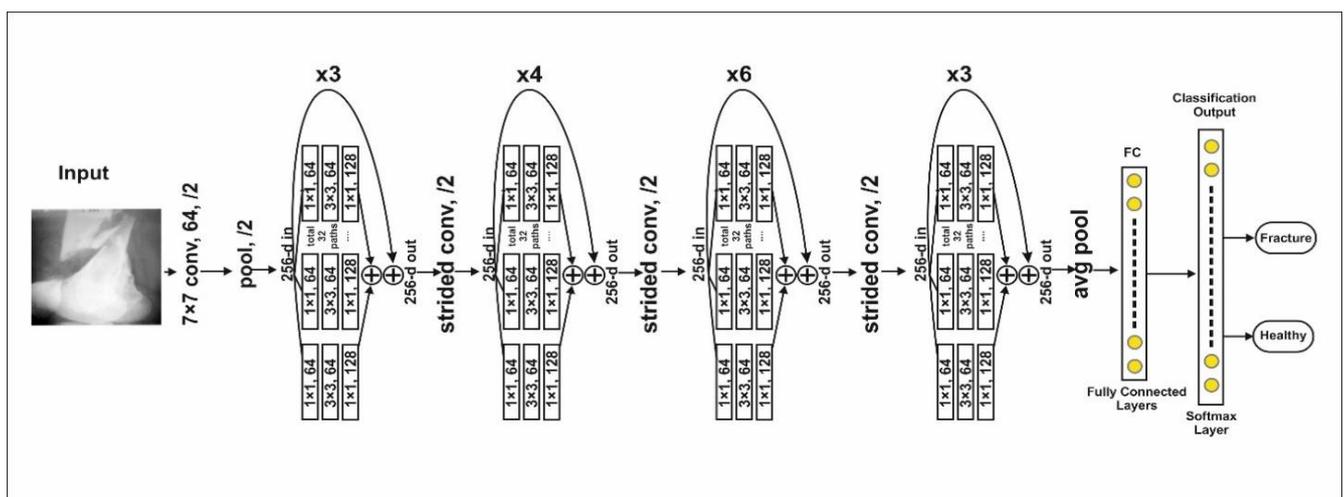


Figure 3. The architecture of the ResNeXt model.

3.4. The VGG16 Model

The network VGG16 [39] is considered one of the superior architectures of the vision model to date. The strangest thing about VGG16 is that it replaces the huge number of hyper parameters. The VGG16 network is 16 layers deep and has a kernel of size 3×3 . In addition to this, five layers of this model have max-pooling layers of 2×2 . After the last max-pooling layer, it has three dense layers, one fully connected layer and Softmax as a

classifier layer. The architecture of the VGG16 for the proposed bone diagnosis is shown in Figure 4.

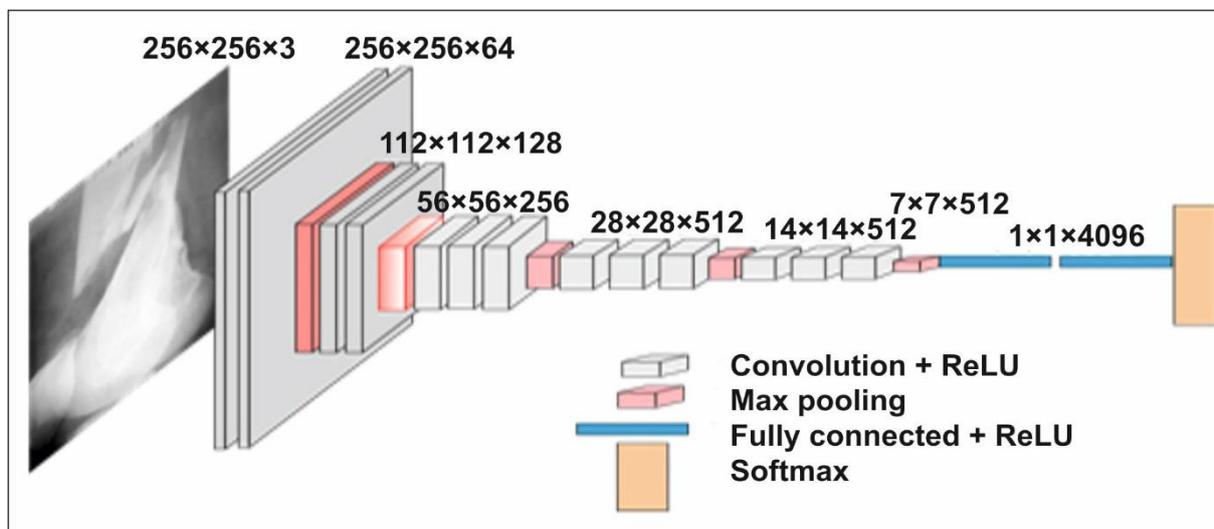


Figure 4. The VGG16 model architecture for bone diagnosis.

The proposed model can be used for real-time human bone diagnosis by creating web and mobile applications to support a doctor. This model has a smaller number of neurons and computation time per epoch is less compared to the other models discussed in the literature, except MobileNetV2. In addition, classification accuracy is much better than the VGG16, AlexNet, ResNeXt and MobileNetV2. The comparison of each model is shown in Table 2.

Table 2. Comparison of different deep CNN model.

Model	Parameters	Time per Epoch	Limitations
VGG16	33×10^6	168 s	This model has a high number of parameters due to long training time
AlexNet	24×10^6	115 s	The performance of the model is not optimal since it is not very deep and it struggles to scan all features.
ResNeXt	23×10^6	140 s	This model is 50 layers deep and requires more training time. Hence, difficult to implement for real-time applications.
MobileNetV2	6.9×10^6	112 s	MobileNet is small in size, small in parameters, and fast in performance. It is less accurate than other state-of-the-art networks.

4. Proposed Hybrid SFNet

In Table 2, we summarized the VGG16, AlexNet, ResNeXt, and MobileNetV2. The VGG16 is a sequential model having 33×10^6 parameters, due to this computation, time is high and overfitting problems may arise. AlexNet is also an eight-layer sequential model with 24×10^6 parameters; this model is less deep due to which it is not able to scan all features. On the other hand, ResNeXt is 50 layers deep with 23×10^6 parameters. Due to this, it has a high training time and a large amount of data are required. MobileNetV2 is 53 layers deep with 6.9×10^6 parameters, this model is lightweight due to its depth-wise convolution architecture, but efficiency is less compared to other state-of-the-art deep CNN models.

In the proposed study, a novel multi-scale hybrid feature fusion model has been developed, which has 5×10^6 parameters with less computation time per epoch compared to other state-of-the-art models discussed above. The complete architecture of the hybrid SFNet is shown in Figure 5. This model contains a two-scale feature extraction model. In

each scale, a six-layer convolution module has been utilized. The size of the first, second, third, fourth, fifth and sixth convolution layers is 16, 32, 64, 128, 256 and 512, respectively. Each of the six layers of the feature extraction module contain a rectified linear unit (ReLU) activation, a batch normalization and a global average-pooling layer with a pool size of 2×2 . The feature extracted from each scale is fused to form a single feature tensor at the concatenation layer. Furthermore, for the classification of bone image, fused tensor is passed to the flattening layer followed by a dense layer of 1024 neurons, ReLU layer, the dropout and the dense layer of two neurons. The classification of bone into healthy and fracture is performed using the Softmax activation function. The Softmax optimizer converts logits into probability. In the present study, feature set x and input vector are passed to the system. The class value is determined by setting the value k to 2 and class labeling is performed using variable j . After that, in each iteration a bias W_0X_0 is added. The mathematical equation is shown in Equation (2).

$$p(y = j|\phi^{(i)}) = \frac{e^{\phi^{(i)}}}{\sum_{j=0}^k e^{\phi_k^{(i)}}} \quad (2)$$

where,

$$\phi = W_0X_0 + W_1X_1 + \dots + W_kX_k \quad (3)$$

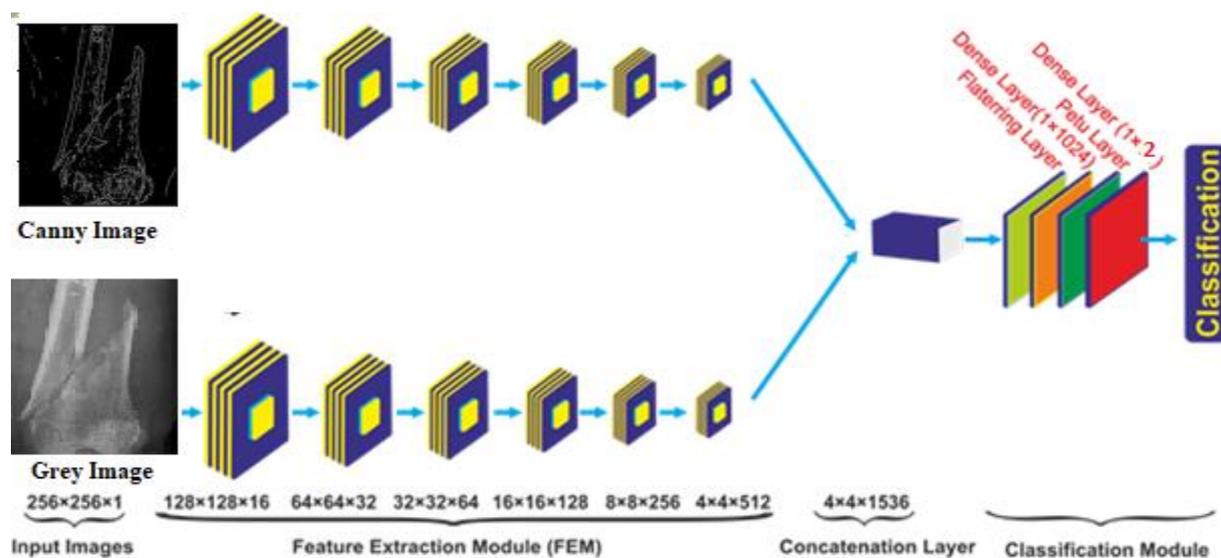


Figure 5. Proposed multi-scale feature fusion CNN model.

4.1. Local Response Normalization

Local Response Normalization (LRN) improves the ability of excited neurons to suppress adjacent neurons. ReLU neurons have unlimited activation and require an LRN that recognizes high frequency function with a large response. Normalizing the local neighborhood of an excited neuron makes it even more sensitive than that neighborhood. The activity of neurons $\alpha_{u,v}^j$ in a place (u, v) by applying the kernel j to the generalization of resources. Then, the non-linearity is applied by ReLU. The $a_{u,v}^j$ is the normalized response, which is calculated as per the Equation (4).

$$b_{u,v}^j = \alpha_{u,v}^j / \left(t + \alpha \sum_{i=\max(0,j,n/2)}^{\min(N,1,j+n/2)} (\alpha_{u,v}^i)^2 \right)^\beta \quad (4)$$

where, N = Total layers, t, α, n, β = different hyper-parameters.

The deep CNN model performance depends on the architecture of the models. Therefore, while designing a model, hyper parameters play an important role. The hyper-parameters t, α, n, β are used to enhance the feature map from one layer to another by finding the intensity of the maximum pixels during the local response normalization process. The division by zero is avoided by setting $t = 2$. The consecutive pixels of input layers that undergo normalization are defined by the n . During the normalization process α is set to 10^{-4} and the contrasting constant β is set to 0.75.

4.2. Deep Feature Fusion

The feature fusion of two-scale SFNet is defined as follows.

$$A = \{a_1, a_2, \dots, a_n\} \quad (5)$$

$$V = \{v_1, v_2, \dots, v_m\} \quad (6)$$

where, $n = 512$ and $m = 512$, after that concatenation of these features is calculated as

$$F_{con} = A \oplus V = (a_1, a_2, \dots, a_n, v_1, v_2, \dots, v_m) \quad (7)$$

where, F_{con} represents the fused feature vector.

4.3. Improved Canny Edge Detection Algorithm

The morphological characteristics of fractured bone and healthy bone are different at the fracture location. In addition, the pixel intensity of the fracture zone is less compared to the healthy zone. Therefore, the localization of these pixels needs a method which can detect edges with fewer pixels value suppression. The traditional canny edge detection is widely used for edge extraction in a digital image [40]. However, the traditional method produces isolated edge points due to Gaussian filtering. In addition, it cannot satisfy algorithm adaptability as threshold values are fixed. Furthermore, it also increases the pseudo edge points. It is necessary to address this problem for fracture bone diagnosis, since the pixel intensities at the fracture points are less. Therefore, some changes are made to the traditional canny edge detection algorithm. First, the Gaussian filter is replaced with a median filter. For a particular location of bone, an image pixel value is replaced with a median value. The median filter for a grey image of $f(x, y)$ with filtering window $W = M_{m \times n}$ is defined as

$$C(x, y) = \text{Med}\{f(x + m, y + n) | M_{m \times n} = 1, (x, y) \in N\} \quad (8)$$

The adaptive threshold is calculated using the difference of adjacent gradient magnitude, as shown below.

$$G_1(i + 1) - G_1(i) \quad (9)$$

Consider an image with pixel intensity $I(x, y)$

$$I(x, y) = \begin{bmatrix} a_0 a_3 a_6 \\ a_1 a_4 a_7 \\ a_2 a_5 a_8 \end{bmatrix} \quad (10)$$

The gradient in x and y direction is calculated as:

$$G_x = (a_6 + 2a_7 + a_8) - (a_0 + 2a_3 + a_6) \quad (11)$$

$$G_y = (a_6 + 2a_7 + a_8) - (a_0 + a_1 + a_2) \quad (12)$$

The magnitude and direction of the central pixel is calculated as:

$$G_1(i, j) = \sqrt{G_x^2 + G_y^2} \quad (13)$$

$$\theta(i, j) = \arctan\left(\frac{G_x}{G_y}\right) \quad (14)$$

Samples of bone fracture image is shown in Figure 6a,c and their corresponding canny image in Figure 6b,d, respectively, as shown in Figure 6.

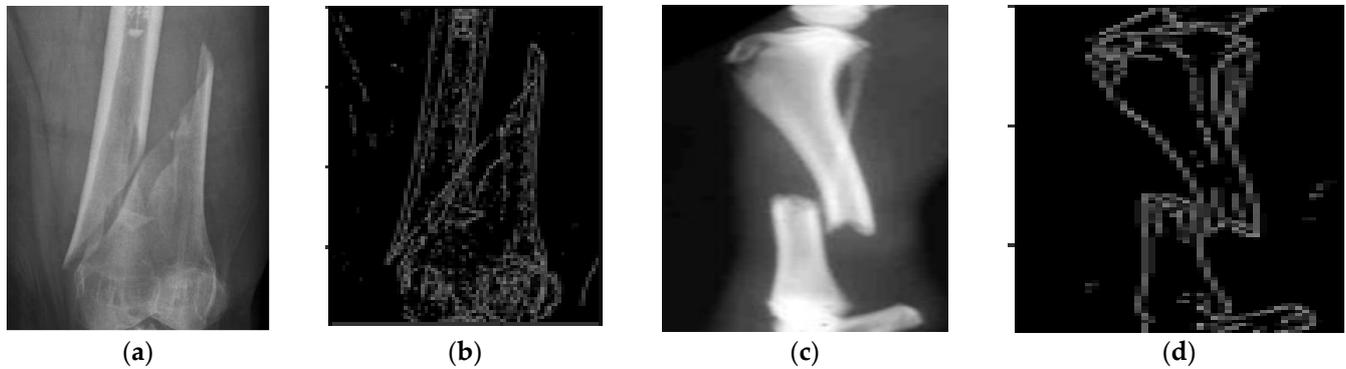


Figure 6. Samples of bone fracture (a,c) and their corresponding canny (b,d), respectively.

The modified canny edge detection algorithm is able to extract the fractured region. The canny image along with the grey image used for training of the model has produced significant improvement in the classification accuracy. The noise can affect the performance of the deep CNN model. Hence, to obtain a highly robust model, noise effect should be minimized. The improved canny edge algorithm can suppress the noise and produce fine edges of the bone image. That helps the proposed model to learn more accurate features from bone images. Proposed method is depicted in Figure 7.

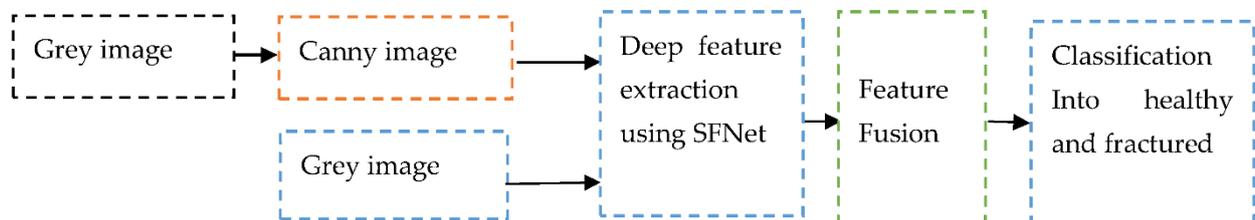


Figure 7. The data flow diagram of the proposed method using hybrid SFNet.

4.4. Loss Function

Loss measures the performance of the model. In order to develop a highly sensitive model, loss should be minimum. In the proposed study, the loss for healthy and fractured bone is calculated as follows.

$$\text{Loss} = \frac{1}{N} \sum_{j=1}^N -((y_j \times \log(P_j)) + (1 - y_j) \times \log(1 - P_j)) \quad (15)$$

where, $N = 2$, $P_j = j$ th scalar value in the model output and $y_j =$ corresponding target value.

5. Results Analysis and Discussion

The results obtained using AlexNet, VGG16, ResNeXt, MobileNetV2 and the proposed hybrid SFNet are discussed in this section. In addition, performance measures and their evaluation methods are also explored with the confusion matrices obtained by these models.

5.1. Dataset

The Bone X-ray image datasets are obtained from a variety of publicly available sources, such as the Cancer Imaging Archive, The Diagnostic Imaging Dataset (DID) [41,42] and Technology, Shibpur (IEST) [43].

5.2. Performance Parameters

The performance measures of the models are calculated using precision, recall, F1-score and accuracy. These parameters are evaluated from the confusion matrix of each model using the mathematical formula shown in Table 3.

Table 3. The mathematical formula to calculate performance measures.

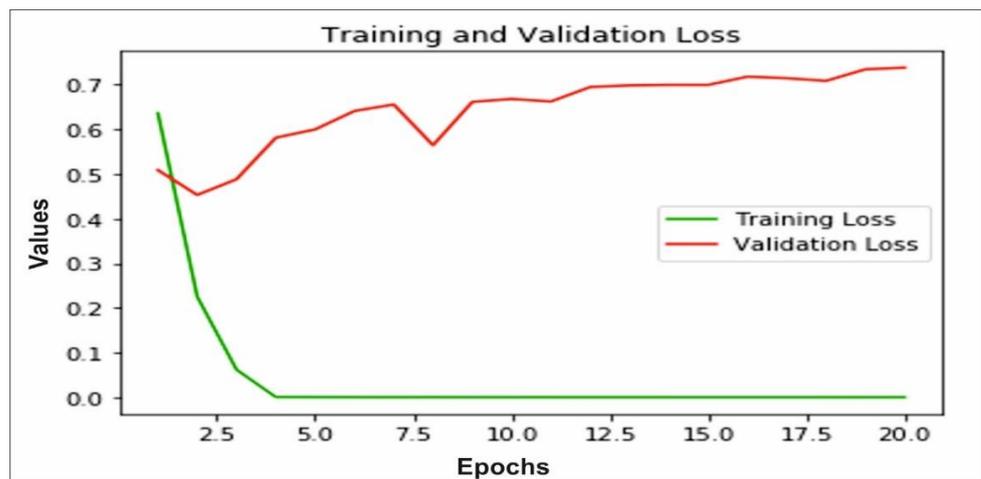
Measures	Formula	Definition
Accuracy	$\text{Accuracy} = \left(\frac{TP+TN}{TP+TN+FP+FN} \right)$	It is calculated by the ratio of the total number of correctly predicted to the total number of test images.
Precision	$\text{Precision} = \left(\frac{TP}{TP+FP} \right)$	The precision is calculated using actual results divided by the total number of true positive samples.
Recall	$\text{Recall} = \left(\frac{TP}{TP+FN} \right)$	The recall is calculated using the total number of positive samples relative to the total number of predictions.
F1-score	$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	The F1-score measures the harmonic mean of the model performance.

The proposed method is implemented using Python 3.6, Tensorflow2, on TITAN X GPU Nvidia GeForce GTX with 128GB RAM, dual graphics card of 8 GB and window 10 operating system. The initial learning rate 3×10^{-3} batch size is 32 and 20 epochs is set for the training of the model.

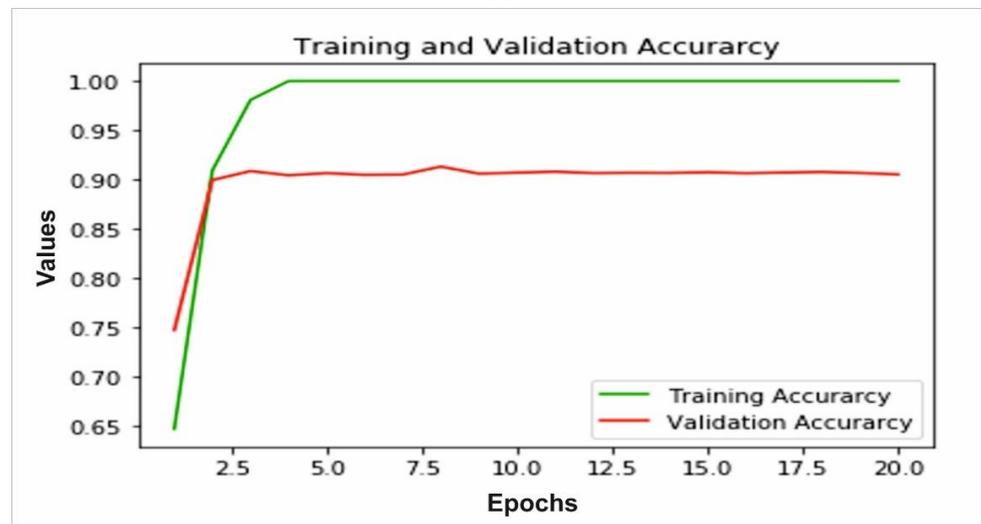
5.3. Result Analysis

The data augmentation technique is also performed to improve the size of the dataset [44]. The basic rotation, flip horizontal, flip vertical and scaling are applied. After data augmentation, the dataset contains 34,000 bone images. Finally, the dataset is randomly divided into training and validation using 80% and 20% ratios, respectively. In this way, each class contain 13,600 images for training and 3400 images for testing. We have maintained the same size $256 \times 256 \times 3$ images for each of the models during training and testing. Each model is trained for 20 epochs with a batch of 32 images. In addition, Adam and RMS Prop Optimizer work with a 3×10^{-3} learning rate for the training of models.

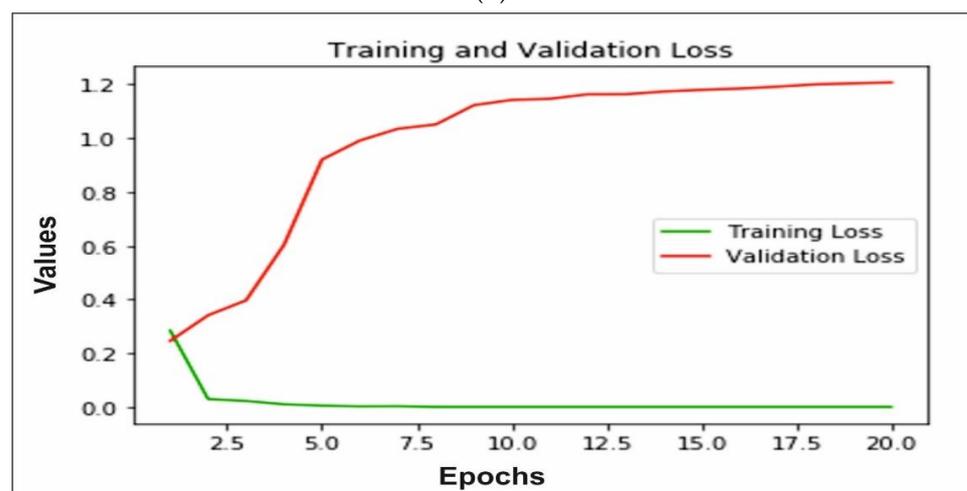
The accuracy and loss of training and validation for each model is calculated to measure the performance. In Figure 8a,b we can observe that for the AlexNet, training loss is very small, whereas the testing loss is increasing; also training accuracy is close to 99% after 10 epochs and testing accuracy decreases after 15 epochs and reaches 90%. A similar pattern can be observed in Figure 8c,d of the VGG16 model. The training loss is very close to 0 and testing loss increases after 10 epochs to reach 1.2. The training accuracy reached 99% and testing accuracy remain close to 94% after 10 epochs. The ResNeXt model loss and accuracy are shown in Figure 8e,f, respectively. The validation and training loss is nearby 0 after 8 epochs and training accuracy reaches 100%, whereas testing accuracy reached 95%. Furthermore, MobileNetV2 trained on ImageNet is implemented for bone diagnosis. In Figure 8g, the training and validation loss of the transfer learning-based approach is depicted. We can observe that the training loss is close to 0, whereas the testing loss increases and reaches 1.2. In addition, the accuracy of training and validation is shown in Figure 8h; the training accuracy reached 99% after 10 epochs and testing accuracy decreases from 70% to 52%. Finally, in Figure 8i,j we can see, using the proposed multi-scale deep learning model, the training and testing loss reached very close to 0, and accuracy reached more than 99%.



(a)

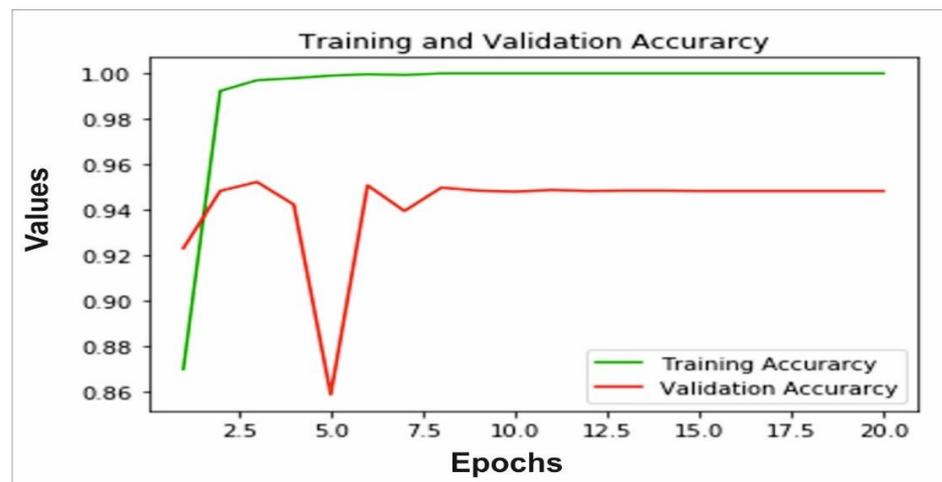


(b)

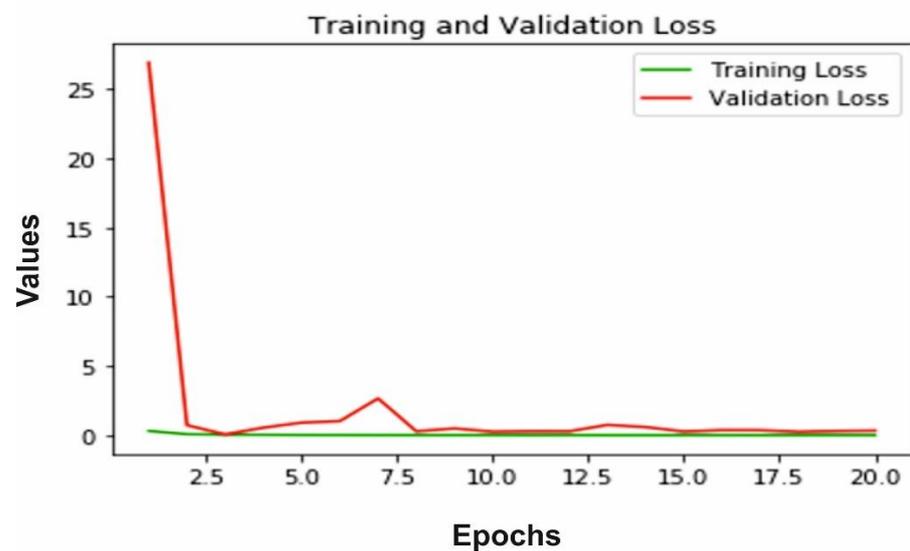


(c)

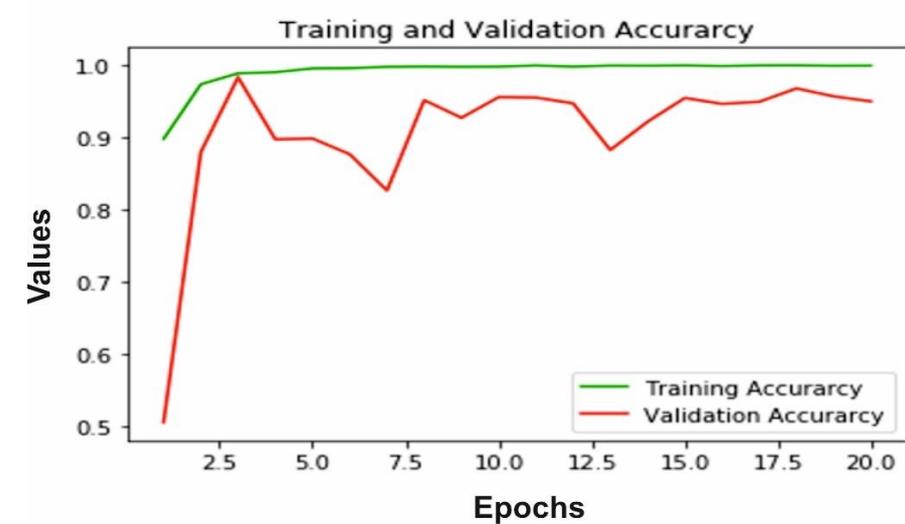
Figure 8. Cont.



(d)

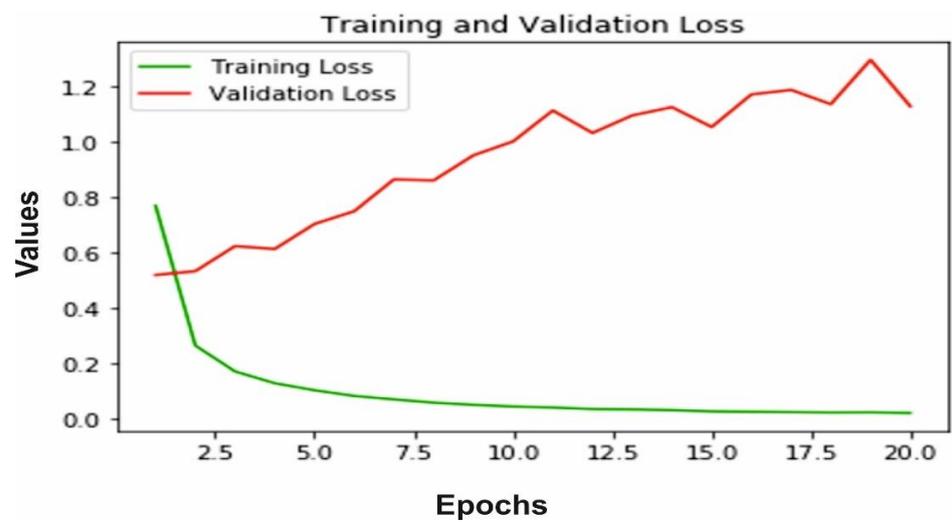


(e)

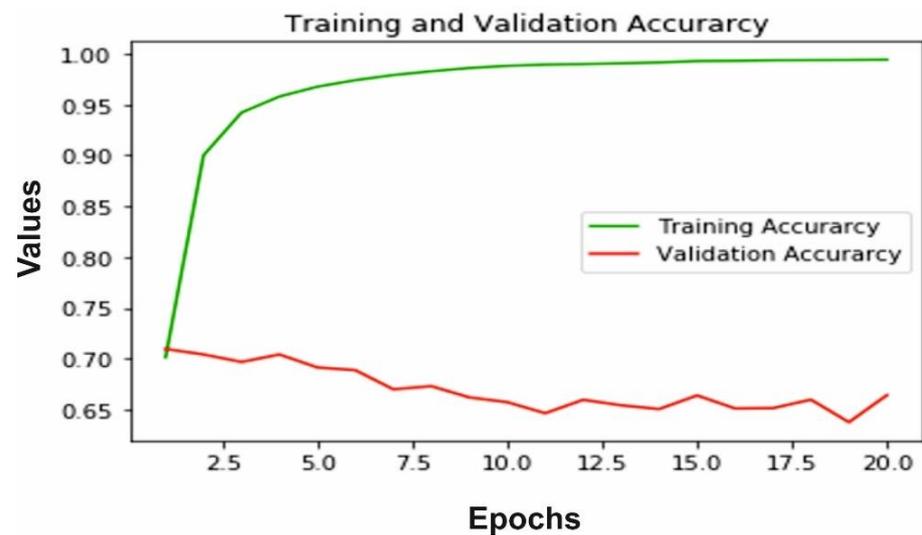


(f)

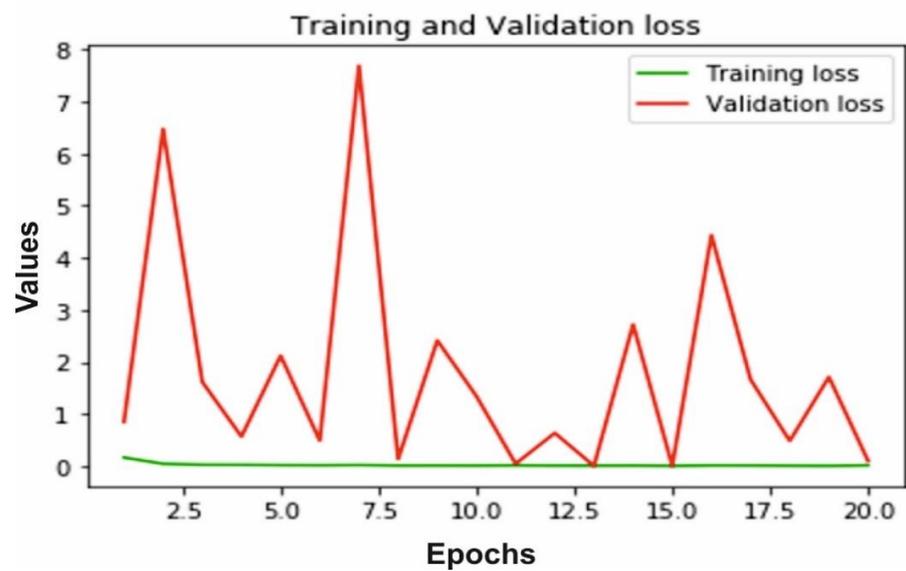
Figure 8. Cont.



(g)



(h)



(i)

Figure 8. Cont.

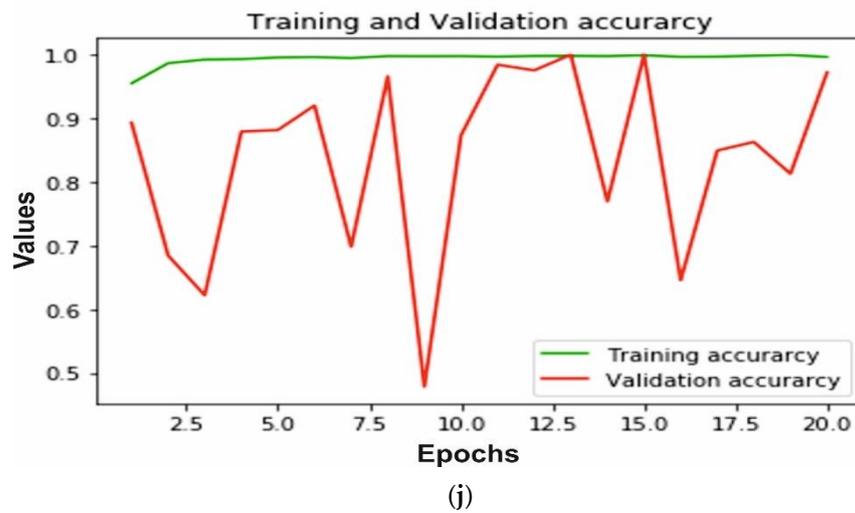


Figure 8. Illustration of training and validation loss of AlexNet, VGG16, ResNeXt, MobileNetV2 and proposed model through (a,c,e,g,i), respectively, and training and validation loss of AlexNet, VGG16, ResNeXt, MobileNetV2 and proposed model through (b,d,f,h,j), respectively.

The confusion matrix for the AlexNet, VGG16, ResNeXt, MobileNetV2 and the proposed hybrid SFNet is shown in the Figure 9a, Figure 9b, Figure 9c, Figure 9d,e, respectively. The false positive value for the AlexNet, VGG16 and ResNeXt model is 0. However, the false negative value of the AlexNet, VGG16 and ResNeXt is 645, 325 and 341, respectively. The pre-trained MobileNetV2 has the highest false positive and false negative values of 2722 and 830, respectively. In contrast to this, the proposed multimodal has the least number of false negative values of 60. The performance measures of the models are measured using precision, recall, F1-score and accuracy.

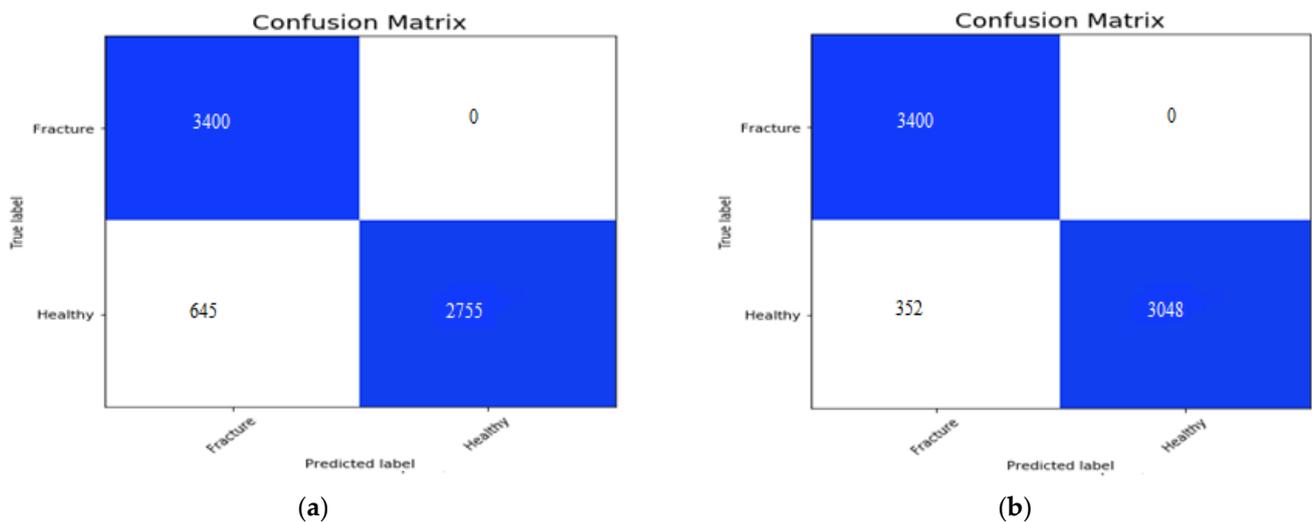


Figure 9. Cont.

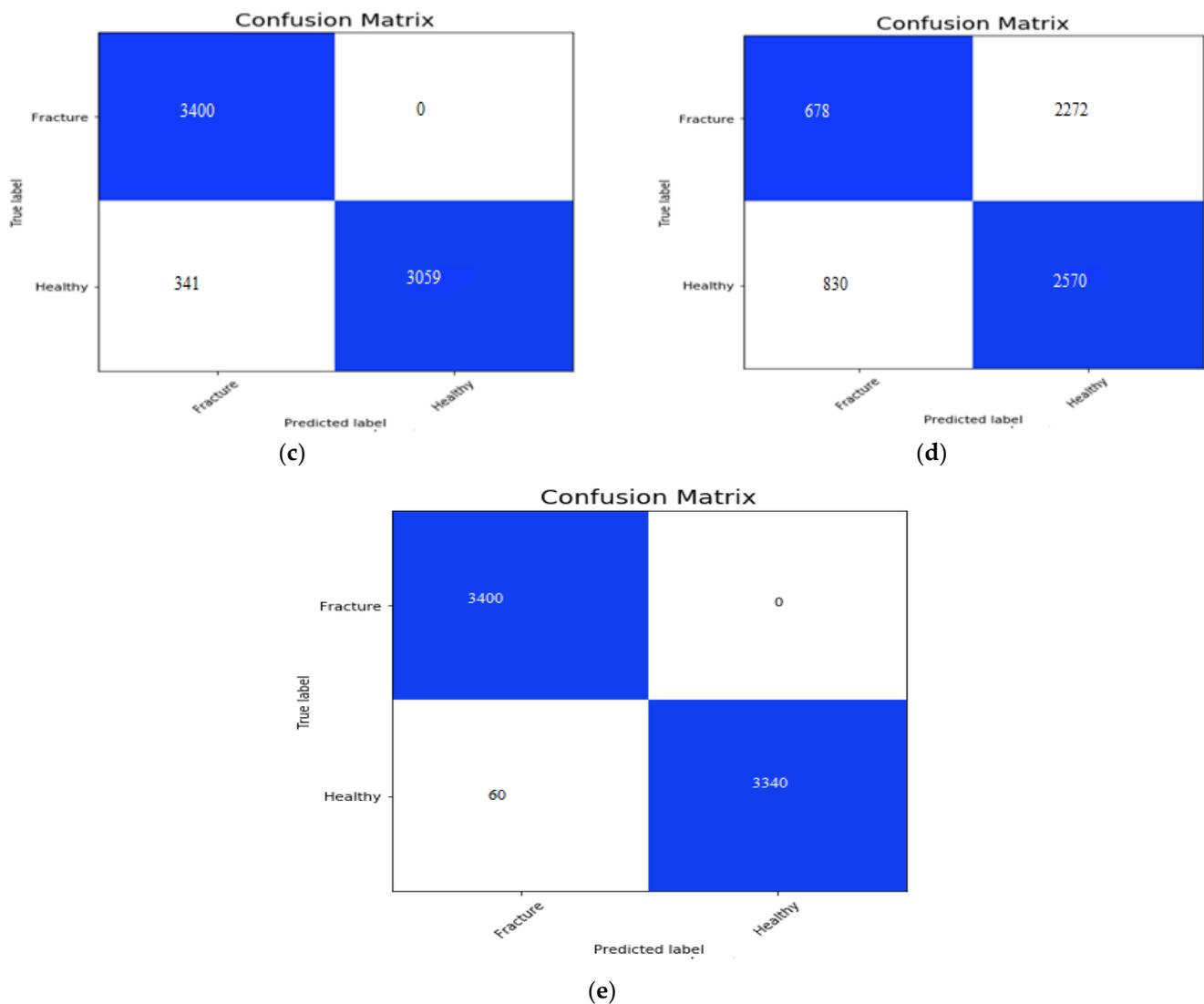


Figure 9. The confusion matrix of the AlexNet, VGG16, ResNeXt, MobileNetV2 and proposed model is shown in the Figures (a–e), respectively.

5.4. Discussion

The performance measures are shown in Table 4 and confusion matrices obtained in Figure 9. In Figure 9a, the confusion matrix of AlexNet has 3400 TP, 2755 TN, 0 FP and 645 FN values. The confusion matrix of VGG16 is shown in Figure 9b, which has 3400 TP, 3048 TN, 0 FP and 352 FN values. In Figure 9c, we can observe 3400 TP, 3059 TN, 0 FP and 341 FN values yielded by ResNeXt. Furthermore, Figure 9d, shows the confusion matrix of pre-trained MobileNetV2, which has 678 TP, 2570 TN, 2272 FP and 830 FN values. Finally, in Figure 9e the confusion matrix obtained by hybrid SFNet is shown, which has 3400 TP, 3340 TN, 0 FP and 60 FN values.

In the Table 4, we can see that the precision of the AlexNet, VGG16, ResNeXt and the proposed hybrid SFNet is 100% for the healthy bone. The least value of precision, 49% for the healthy bone, can be seen in the MobileNetV2. Similarly, for fractured bone, the least precision value of 45% was achieved using MobileNetV2 and the highest value of 100% was observed using the proposed hybrid SFNet. The recall value of 100% was observed in fractured bone using AlexNet, VGG16 and ResNeXt, whereas, the proposed model achieved 98% and the least value of 20% was shown by MobileNetV2. In addition, an F1-score of 99% was achieved by the hybrid SFNet, whereas other models reached close to 95%. These results conclude that the performance of the pre-trained MobileNetV2 is not

optimal; however, AlexNet, VGG16 and ResNeXt trained on bone image performance is outstanding, but require high computation time and further improvement for fractured bone diagnosis. The proposed hybrid SFNet has shown remarkable performance with the least computation time.

Table 4. The comparison of performance measures.

Model	Types of Bone	Precision	Recall	F1-Score	Accuracy
AlexNet	Fracture	0.84	1	0.91	0.91
	Healthy	1	0.81	0.90	
VGG 16	Fracture	0.91	1	0.95	0.95
	Healthy	1	0.90	0.95	
ResNeXt	Fracture	0.91	1	0.95	0.95
	Healthy	1	0.90	0.95	
MobileNetV2	Fracture	0.45	0.20	0.28	0.48
	Healthy	0.49	0.76	0.59	
Proposed hybrid SFNet	Fracture	1	0.98	0.99	0.99
	Healthy	0.98	1	0.99	

In addition to this, the recall value of the AlexNet, VGG16, ResNeXt and the proposed model is 100%. However, the highest precision value of 100% is achieved for the fractured bone by the proposed model. Furthermore, the proposed hybrid SFNet model has a classification accuracy of 99% and the other models, such as AlexNet, VGG16, ResNeXt and MobileNetV2, have 91%, 95%, 95% and 48%, respectively.

6. Comparative Analysis

A comparative study of the deep learning-based convolution neural network (DCNN) model has been demonstrated. In the past research, some methods are available for the long bone fracture diagnosis. These methods are supportive to a doctor in their diagnosis process, but they are less efficient. Hence, an efficient and robust model for bone diagnosis can be developed, Haitaamar et al. [17] applied U-Net for the segmentation of the bone image. After segmentation, classification is performed with an accuracy of 95%. Nguyen et al. [18] applied YoloV4 for the real-time bone fracture diagnosis and achieved an accuracy of 81.19%. Similarly, Wang et al. [19] developed DCNN for the diagnosis of bone images. Their method confirmed 88.7% classification accuracy.

Wang et al. [20] developed a two-stage deep CNN based method for bone fracture diagnosis and achieved an accuracy of 87.8%. Yahalomi et al. [21] and Abbas et al. [22] applied faster-RCNN for bone diagnosis and achieved an accuracy of 96% and 97%, respectively. Sasidhar et al. [45] evaluated the performance of VGG19, DenseNet121, and DenseNet169. The highest classification accuracy of 92% using DenseNet169 is achieved. All of the methods discussed above are capable of bone diagnosis. However, these methods' performance is not optimal and real-time application is not available. In the proposed study, we have compared the state-of-the-art deep CNN models AlexNet, VGG16, ResNeXt and MobileNetV2. The performance of the model is evaluated on a publicly available dataset.

The training dataset size plays an important role to avoid model overfitting. To avoid the overfitting problems, we need to trained the deep CNN model on a large volume dataset. One way to improve the size of the dataset is data augmentation. Therefore, in this research, a data augmentation technique has been applied.

After data augmentation, the size of the augmented dataset is 34,000. The healthy and fractured bone image contains an equal number of 17,000 images and the dataset is divided randomly into 80% for the training and 20% for the validation. None of the testing images have been included in the training of the model. All simulations are done using Keras library with back-end TensorFlow on window 10 operating system with 128 GB RAM with dual 8 GB graphics. We have taken care that the size of the image will be 256×256 for each model used in this study. In addition, each model is trained for

20 epochs with a batch of 32 images. The Adam optimizer has been used for AlexNet, ResNeXt, MobileNetV2, and the proposed multimodal deep learning model with an initial learning rate of 3×10^{-3} . In the experiment, we found that the Adam optimizer does not produce an optimal performance with the VGG16; flipping to the RMSprop with initial learning of 0.00001 produces a satisfactory performance for the bone diagnosis.

In addition, Table 5 summarizes the different state-of-the-art methods for bone fraction detection. We can see in Table 5 that hybrid SFNet with only grey image and faster R-CNN have the same classification accuracy of 97%. However, SFNet with grey and canny image performance is highest compared to the state-of-the-art model. This high performance is achieved by the hybrid SFNet model using canny image, due to the localization of fracture region as described in Algorithm 1 and Section 4.3. The features learned by the model from the localized region improves the training and classification accuracy.

Algorithm 1: Bone diagnosis technique using hybrid SFNet

- 1: Create a fractured and healthy image dataset.
 - 2: Apply augmentation technique rotation, flip horizontal, flip vertical and scaling to increase the size of the dataset.
 - 3: Find the edge in an image using the improved canny edge algorithm discussed in Section 4.3.
 - 4: For $I = 1$ to 20 train the model
 - (a): Input grey and canny images to hybrid SFNet
 - (b): Apply Equations (9) and (10) to convert logits into probability values
 - (c): Calculate training and validation loss for each epoch using equation 155: Find overall training accuracy using the equation discussed in Table 3.
 - 6: Find overall validation accuracy using the equation discussed in Table 3.
 - 7: Find the loss of the hybrid SFNet.
 - 8: Plot a training and validation loss graph for 20 epochs.
-

Table 5. Comparison table based on accuracy of the models.

Study	Model	Accuracy
Haitaamar et al. [17]	U-Net	95%
Nguyen et al. [18]	YOLOv4	81.91%
Wang et al. [19]	DCNN	88.7%
Ma et al. [20]	Faster R-CNN	90.11
Wang et al. [21]	Two-stage R-CNN	87.8%
Yahalomi et al. [22]	Faster R-CNN	96%
Abbas et al. [23]	Faster R-CNN	97%
Sasidhar et al. [45]	VGG19, DenseNet121, DenseNet169	92%
Proposed method	Hybrid SFNet + Grey	97%
	Hybrid SFNet + Canny + Grey	99.12%

The performance measures precision for each model is shown in Figure 10. In Figure 10, it can be noticed that the highest precision value of 100% for fractured bone is obtained by the hybrid SFNet. The VGG16 and ResNeXt have a similar value of 91%. Whereas, AlexNet showed 84% and the least value of 45% is yielded by the pre-trained MobileNetV2.

In Figure 11, the recall value of each model is plotted. The recall value of AlexNet, VGG16 and ResNeXt for fractured bone is 100%. Whereas, Hybrid SFNet achieved 98% and MobileNetV2 obtained 20% recall values. For healthy bone, hybrid SFNet has 100% recall value and other models reached 90%, except MobileNetV2.

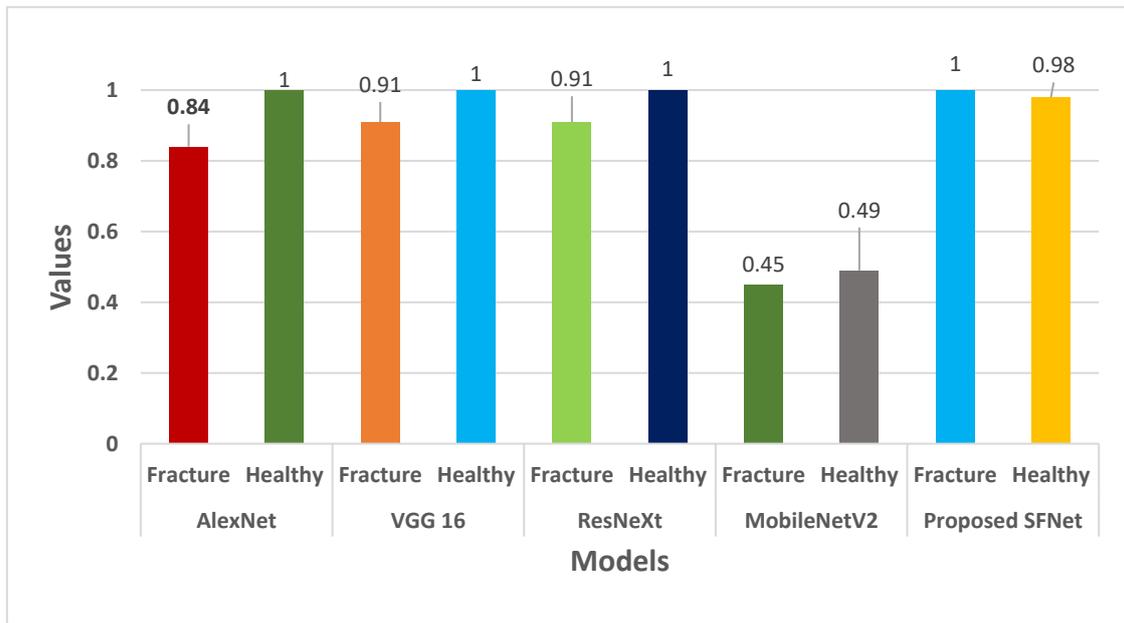


Figure 10. Precision comparison of the state-of-the-art method with hybrid SFNet.

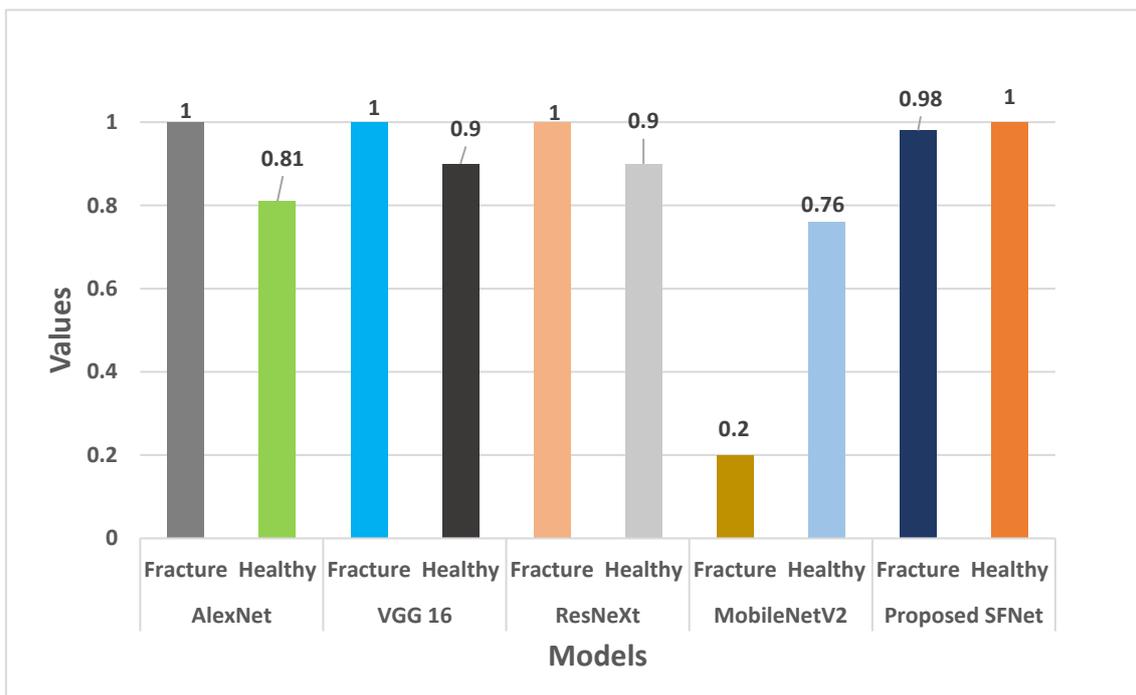


Figure 11. Recall comparison of the state-of-the-art method with hybrid SFNet.

Furthermore, the F-score for each model is shown in Figure 12. For both fractured and healthy bone, the proposed SFNet achieved a 99% F1-score. Whereas, VGG16 and ResNeXt obtained 95% and AlexNet a 91% F1-score. The least F1-score of 59% for healthy bone was observed in MobileNetV2.

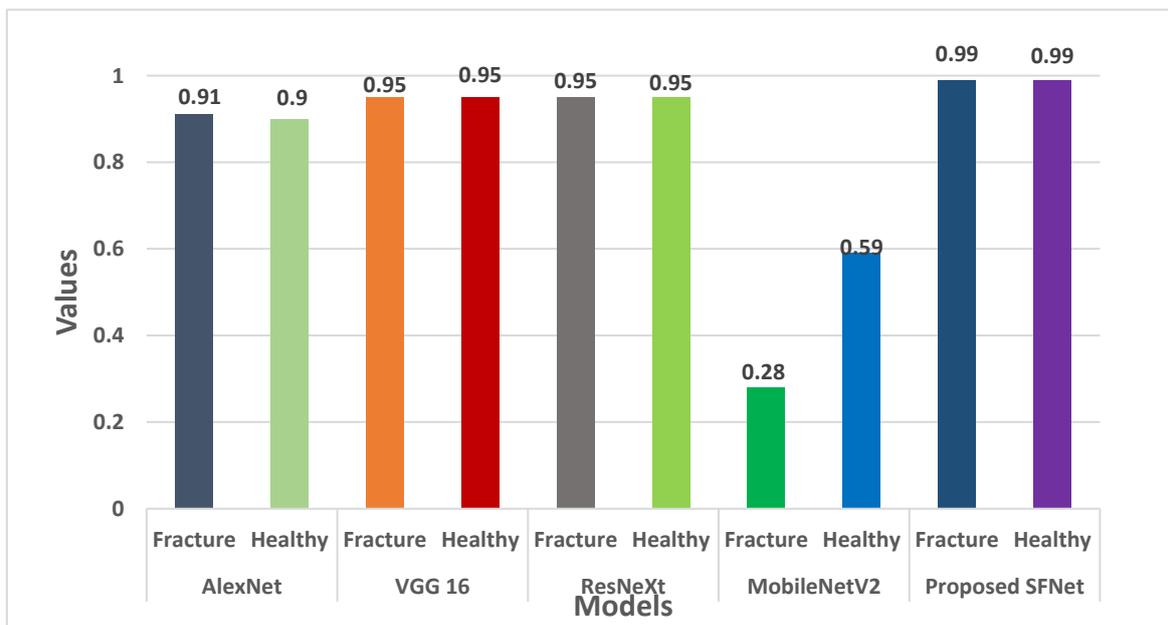


Figure 12. F1-score comparison of the state-of-the-art method with hybrid SFNet.

Finally, classification accuracy of each model is shown in Figure 13. In Figure 13, we can notice that the highest classification of more than 99% is obtained by hybrid SFNet. The VGG16 and ResNeXt achieved remarkable accuracy of 95%. The AlexNet and MobileNetV2 performance is not optimal.

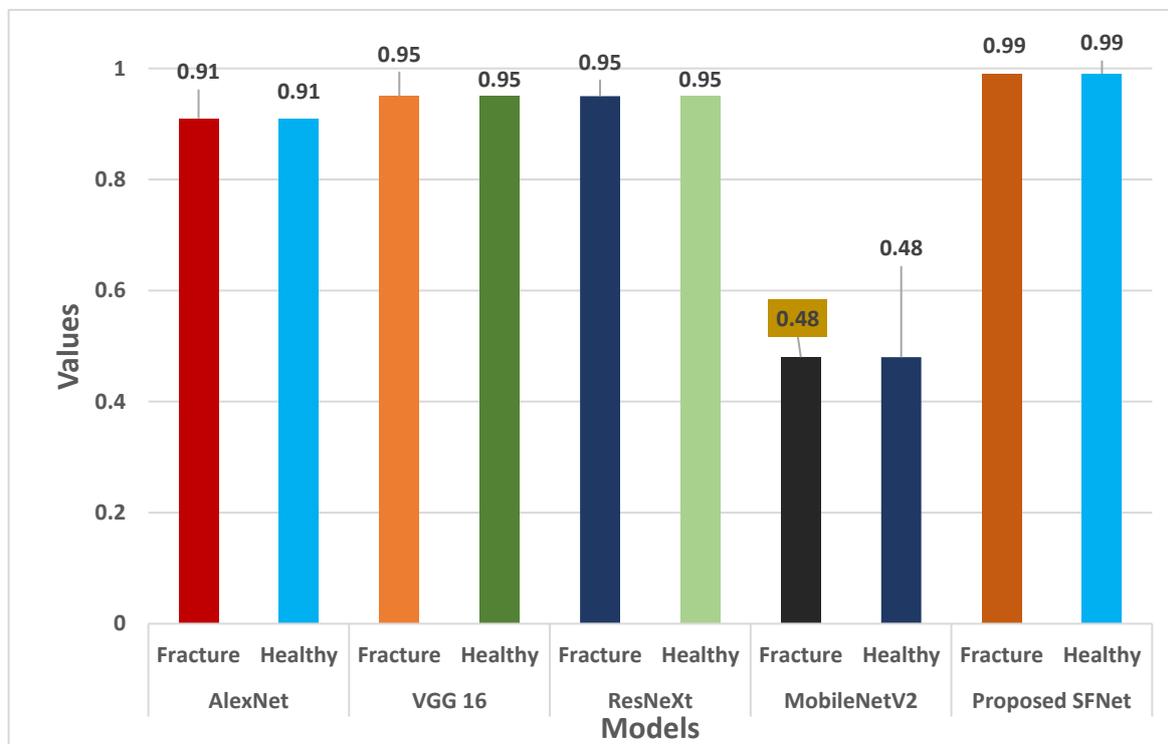


Figure 13. Accuracy comparison of the state-of-the-art method with hybrid SFNet.

We can conclude from Figures 10–13 that the performance measures of the proposed SFNet is highest as compared to the state-of-the-art deep CNN model. These performance

measures indicate that the proposed hybrid SFNet deep learning model are supportive to a doctor to take a second opinion on the fracture and healthy bone diagnosis.

7. Conclusions and Future Work

In the present work, we have proposed a novel multimodal feature fusion hybrid SFNet. The SFNet has less parameters compared to AlexNet, VGG16 and ResNeXt. In addition, it takes less computation time per epochs, as discussed in Table 2. The grey image and their corresponding canny image are fed to the hybrid SFNet for deep feature extraction. After feature extraction, a pool of features is passed to the classification module. The classification of healthy and fractured images is performed using the Softmax layer. Furthermore, the performance measures by F1-score, precision and recall for the fractured bone as well as healthy bone are calculated. The achieved accuracy of 99.12% and precision of 99% by the hybrid SFNet indicates that canny image improves the performance. Furthermore, the training and validation losses are close to 0. This confirms the model is highly sensitive.

In future studies, we will locate the size of the fracture region, which will help experts to take more action related to bone fracture. In addition, algorithm computation cost will be reduced using optimization techniques. Cloud computing and machine learning play a crucial role in illness detection, but mostly among those who live in distant places with few medical services. While cloud-based systems can enable telehealth services and remote diagnostics, diagnosis tools based on machine learning operate as secondary readers and help radiologists in the accurate diagnosis of illnesses [46]. Due to its capacity to deal with uncertainty and imprecision, neutrosophic theory is frequently employed to solve challenges in medical image processing. Due to the presence of fuzziness and the constrained subjectivity of the specialists utilizing medical pictures, diagnosis is one of the challenging jobs [47]. Image segmentation is one of the most difficult procedures due to specialists' limited observation and uncertainty in medical understanding. Because imprecise information is commonly employed in decision making, crisp values are insufficient to simulate real-world situations. The intuitive set is a helpful tool for describing images with unclear information [48,49].

Author Contributions: Conceptualization, A.S. and B.S.; data curation, D.P.Y.; formal analysis, A.S., S.A., A.B., B.S. and I.B.D.; investigation, S.A. and I.B.D.; methodology, D.P.Y., A.S. and B.S.; project administration, S.A., A.B. and I.B.D.; resources, I.B.D.; software, S.A. and A.B.; validation, B.S.; visualization, I.B.D.; writing—original draft, D.P.Y., A.B. and B.S.; writing—review and editing, I.B.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: There are no available data to be stated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Amirkolaee, H.A.; Bokov, D.O.; Sharma, H. Development of a GAN architecture based on integrating global and local information for paired and unpaired medical image translation. *Expert Syst. Appl.* **2022**, *203*, 117421. [[CrossRef](#)]
2. Singh, L.K.; Garg, H.; Khanna, M. Deep learning system applicability for rapid glaucoma prediction from fundus images across various data sets. *Evol. Syst.* **2022**, 1–30. [[CrossRef](#)]
3. Garg, H.; Gupta, N.; Agrawal, R.; Shivani, S.; Sharma, B. A real-time cloud-based framework for glaucoma screening using EfficientNet. *Multimed. Tools Appl.* **2022**, 1–22. [[CrossRef](#)]
4. Yu, Y.; Rashidi, M.; Samali, B.; Mohammadi, M.; Nguyen, T.N.; Zhou, X. Crack detection of concrete structures using deep convolutional neural networks optimized by enhanced chicken swarm algorithm. *Struct. Health Monit.* **2022**, 14759217211053546. [[CrossRef](#)]
5. Lu, S.; Wang, S.; Wang, G. Automated universal fractures detection in X-ray images based on deep learning approach. *Multimed. Tools Appl.* **2022**, 1–17. [[CrossRef](#)]

6. Guan, B.; Zhang, G.; Yao, J.; Wang, X.; Wang, M. Arm fracture detection in X-rays based on improved deep convolutional neural network. *Comput. Electr. Eng.* **2020**, *81*, 106530. [[CrossRef](#)]
7. Anu, T.C.; Raman, R. Detection of bone fracture using image processing methods. *Int. J. Comput. Appl.* **2015**, *975*, 8887.
8. Bandyopadhyay, O.; Biswas, A.; Bhattacharya, B.B. Long-bone fracture detection in digital X-ray images based on digital-geometric techniques. *Comput. Methods Programs Biomed.* **2016**, *123*, 2–14. [[CrossRef](#)]
9. Basha, C.Z.; Reddy, M.R.K.; Nikhil, K.H.S.; Venkatesh, P.S.M.; Asish, A.V. Enhanced computer aided bone fracture detection employing X-ray images by Harris Corner technique. In Proceedings of the 2020 IEEE Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 11–13 March 2020; pp. 991–995.
10. Hoang, N.D.; Nguyen, Q.L. A novel method for asphalt pavement crack classification based on image processing and machine learning. *Eng. Comput.* **2019**, *35*, 487–498. [[CrossRef](#)]
11. Lin, X.; Yan, Z.; Kuang, Z.; Zhang, H.; Deng, X.; Yu, L. Fracture R-CNN: An anchor-efficient anti-interference framework for skull fracture detection in CT images. *Med. Phys.* **2022**. [[CrossRef](#)]
12. Kitamura, G.; Chung, C.Y.; Moore, B.E. Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation. *J. Digit. Imaging* **2019**, *32*, 672–677. [[CrossRef](#)] [[PubMed](#)]
13. Choi, J.W.; Cho, Y.J.; Ha, J.Y.; Lee, Y.Y.; Koh, S.Y.; Seo, J.Y.; Choi, Y.H.; Cheon, J.; Phi, J.H.; Kim, I.; et al. Deep learning-assisted diagnosis of pediatric skull fractures on plain radiographs. *Korean J. Radiol.* **2022**, *23*, 343–354. [[CrossRef](#)] [[PubMed](#)]
14. Kitamura, G. Deep learning evaluation of pelvic radiographs for position, hardware presence, and fracture detection. *Eur. J. Radiol.* **2020**, *130*, 109139. [[CrossRef](#)] [[PubMed](#)]
15. Kim, D.H.; MacKinnon, T. Artificial intelligence in fracture detection: Transfer learning from deep convolutional neural networks. *Clin. Radiol.* **2018**, *73*, 439–445. [[CrossRef](#)] [[PubMed](#)]
16. Yang, L.; Gao, S.; Li, P.; Shi, J.; Zhou, F. Recognition and Segmentation of Individual Bone Fragments with a Deep Learning Approach in CT Scans of Complex Intertrochanteric Fractures: A Retrospective Study. *J. Digit. Imaging* **2022**, 1–9. [[CrossRef](#)]
17. Haitamar, Z.N.; Abdulaziz, N. Detection and Semantic Segmentation of Rib Fractures using a Convolutional Neural Network Approach. In Proceedings of the 2021 IEEE Region 10 Symposium (TENSYMP), Jeju-si, Korea, 23–25 August 2021; pp. 1–4.
18. Nguyen, H.P.; Hoang, T.P.; Nguyen, H.H. A deep learning based fracture detection in arm bone X-ray images. In Proceedings of the 2021 IEEE International Conference on Multimedia Analysis and Pattern Recognition (MAPR), Hanoi, Vietnam, 15–16 October 2021; pp. 1–6.
19. Wang, M.; Zhang, G.; Guan, B.; Xia, M.; Wang, X. Multiple Reception Field Network with Attention Module on Bone Fracture Detection Task. In Proceedings of the 2021 IEEE 40th Chinese Control Conference (CCC), Shanghai, China, 26–28 July 2021; pp. 7998–8003.
20. Ma, Y.; Luo, Y. Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network. *Inform. Med. Unlocked* **2021**, *22*, 100452. [[CrossRef](#)]
21. Wang, M.; Yao, J.; Zhang, G.; Guan, B.; Wang, X.; Zhang, Y. ParallelNet: Multiple backbone network for detection tasks on thigh bone fracture. *Multimed. Syst.* **2021**, *27*, 1091–1100. [[CrossRef](#)]
22. Yahalomi, E.; Chernofsky, M.; Werman, M. Detection of distal radius fractures trained by a small set of X-ray images and Faster R-CNN. In *Intelligent Computing—Proceedings of the Computing Conference*; Springer: Cham, Switzerland, 2019; pp. 971–981.
23. Abbas, W.; Adnan, S.M.; Javid, M.A.; Ahmad, W.; Ali, F. Analysis of tibia-fibula bone fracture using deep learning technique from X-ray images. *Int. J. Multiscale Comput. Eng.* **2021**, *19*, 25–39. [[CrossRef](#)]
24. Luo, J.; Kitamura, G.; Doganay, E.; Arefan, D.; Wu, S. Medical knowledge-guided deep curriculum learning for elbow fracture diagnosis from X-ray images. In *Medical Imaging 2021: Computer-Aided Diagnosis*; International Society for Optics and Photonics: Bellingham, WA, USA, 2021; Volume 11597, p. 1159712.
25. Beyaz, S.; Açıcı, K.; Sümer, E. Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches. *Jt. Dis. Relat. Surg.* **2020**, *31*, 175. [[CrossRef](#)]
26. Jones, R.M.; Sharma, A.; Hotchkiss, R.; Sperling, J.W.; Hamburger, J.; Ledig, C.; Lindsey, R.V. Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. *NPJ Digit. Med.* **2020**, *3*, 1–6. [[CrossRef](#)]
27. Dupuis, M.; Delbos, L.; Veil, R.; Adamsbaum, C. External validation of a commercially available deep learning algorithm for fracture detection in children. *Diagn. Interv. Imaging* **2022**, *103*, 151–159. [[CrossRef](#)] [[PubMed](#)]
28. Hardalaç, F.; Uysal, F.; Peker, O.; Çiçeklidağ, M.; Tolunay, T.; Tokgöz, N.; Mert, F. Fracture Detection in Wrist X-ray Images Using Deep Learning-Based Object Detection Models. *Sensors* **2022**, *22*, 1285. [[CrossRef](#)] [[PubMed](#)]
29. Pranata, Y.D.; Wang, K.C.; Wang, J.C.; Idram, I.; Lai, J.Y.; Liu, J.W.; Hsieh, I.H. Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images. *Comput. Methods Programs Biomed.* **2019**, *171*, 27–37. [[CrossRef](#)]
30. Mutasa, S.; Varada, S.; Goel, A.; Wong, T.T.; Rasiej, M.J. Advanced deep learning techniques applied to automated femoral neck fracture detection and classification. *J. Digit. Imaging* **2020**, *33*, 1209–1217. [[CrossRef](#)]
31. Weikert, T.; Noordtzi, L.A.; Bremerich, J.; Stieltjes, B.; Parmar, V.; Cyriac, J.; Sauter, A.W. Assessment of a deep learning algorithm for the detection of rib fractures on whole-body trauma computed tomography. *Korean J. Radiol.* **2020**, *21*, 891. [[CrossRef](#)]
32. Tanzi, L.; Vezzetti, E.; Moreno, R.; Aprato, A.; Audisio, A.; Masse, A. Hierarchical fracture classification of proximal femur X-ray images using a multistage deep learning approach. *Eur. J. Radiol.* **2020**, *133*, 109373. [[CrossRef](#)] [[PubMed](#)]

33. Lotfy, M.; Shubair, R.M.; Navab, N.; Albarqouni, S. Investigation of focal loss in deep learning models for femur fractures classification. In Proceedings of the 2019 IEEE International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, 19–21 November 2019; pp. 1–4.
34. Tanzi, L.; Vezzetti, E.; Moreno, R.; Moos, S. X-ray bone fracture classification using deep learning: A baseline for designing a reliable approach. *Appl. Sci.* **2020**, *10*, 507. [[CrossRef](#)]
35. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Processing Syst.* **2012**, *25*, 1–9. [[CrossRef](#)]
36. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
37. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
38. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
39. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
40. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *6*, 679–698. [[CrossRef](#)]
41. Available online: <https://www.cancerimagingarchive.net/collections/> (accessed on 18 June 2022).
42. Available online: <https://www.england.nhs.uk/statistics/category/statistics/diagnostic-imaging-dataset/> (accessed on 18 June 2022).
43. Available online: <https://www.iiests.ac.in/> (accessed on 18 June 2022).
44. Mikołajczyk, A.; Grochowski, M. Data augmentation for improving deep learning in image classification problem. In Proceedings of the 2018 IEEE International Interdisciplinary PhD Workshop (IIPhDW), Swinoujscie, Poland, 9–12 May 2018; pp. 117–122.
45. Sasidhar, A.; Thanabal, M.S.; Ramya, P. Efficient Transfer learning Model for Humerus Bone Fracture Detection. *Ann. Rom. Soc. Cell Biol.* **2021**, *25*, 3932–3942.
46. Lahoura, V.; Singh, H.; Aggarwal, A.; Sharma, B.; Mohammed, M.A.; Damaševičius, R.; Cengiz, K. Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. *Diagnostics* **2021**, *11*, 241. [[CrossRef](#)] [[PubMed](#)]
47. Koundal, D.; Sharma, B. Challenges and future directions in neutrosophic set-based medical image analysis. In *Neutrosophic Set in Medical Image Analysis*; Academic Press: Cambridge, MA, USA, 2019; pp. 313–343.
48. Koundal, D.; Sharma, B.; Guo, Y. Intuitionistic based segmentation of thyroid nodules in ultrasound images. *Comput. Biol. Med.* **2020**, *121*, 103776. [[CrossRef](#)] [[PubMed](#)]
49. Koundal, D.; Sharma, B.; Gandotra, E. Spatial intuitionistic fuzzy set based image segmentation. *Imaging Med.* **2017**, *9*, 95–101.