Int J Syst Assur Eng Manag (March 2022) 13(Suppl. 1):S658–S672
https://doi.org/10.1007/s13198-021-01580-3

ORIGINAL ARTICLE

BoostNet: a method to enhance the performance of deep learning model on musculoskeletal radiographs X-ray images

Pawan Kumar Mall[1] · **Pradeep Kumar Singh**[1]

Received: 26 October 2021 / Revised: 25 November 2021 / Accepted: 30 November 2021 / Published online: 4 January 2022

Abstract In clinical treatment, deep learning plays a pivotal role in medical image classification. Deep learning techniques provide opportunities for radiologists and orthopedic to ease out their lives with faster and more accurate results. The traditional deep learning approach nevertheless reached its performance ceiling. Therefore, in this paper, we investigate different enhancement techniques to boost the deep neural networks performance and provide a solution as BoostNet. The experiment is categorized into four different phases. We have selected ChampNet from benchmark deep learning models (EfficientNet: B0, MobileNet, ResNet18, VGG19). This phase helps to obtain the best model. In the second phase, The ChampNet evaluates with different resolution datasets. This phase helps to finalize the dataset resolution to enhance the performance of ChampNet. In the third phase, Champ-Net merges with image enhancement techniques, Contrast Limited Adaptive Histogram Equalization (CLAHE), High-frequency filtering (HEF), and Unsharp masking (UM). This phase helps to obtain Boost-Net with enriched performance. The last phase helps us to verify BoostNet results with Lightness Order Error. The presented research work fuses the image enhancement technique with ChampNet to generate BoostNet models. An assessment was performed on the Musculoskeletal Radiograph Bone Classification using classification schemes to demonstrate the proposed model's performance. The Classification accuracy of BoostNet was for the train a test dataset with and without enhancement techniques. The proposed model ChampNet + CLAHE, ChampNet + HEF, ChampNet + UM approach achieved 95.88%, 94.99%, and 94.18% accuracy, respectively. This experiment leads to a more accurate and efficient classification model. The main aim of this paper is to enhance techniques to boost the deep neural networks performance and provide a solution as BoostNet.

Keywords Classification · Deep Learning (DL) · Image Enhancement · Musculoskeletal Radiograph · Medical Image

1 Introduction

In medical imaging, a wide variety of studies have been conducted using different models of deep neural networks (DNNs) to classify or diagnose diseases. Technologies such as DNNs and computer vision have already demonstrated the ability to recognize images and exceed human accuracy. Beside, the remarkable success in recent years of deep learning models (DLMs) in image classification, segmentation, and detection tasks, connects to an era of significant use for diagnostic medical imaging. However, the availability of large datasets with reliable ground-truth analysis is a major issue in medical imaging.

DLMs have contributed to a series of breakthroughs in the task of image classification (Krizhevsky et al. 2009, 2012). The DLMs utilities low, mid, and high-level features (Zeiler and Fergus 2014). Bundy A., & Wallen, L. et al. shows that the depth of a network model is of critical

✉ Pawan Kumar Mall
pawankumar.mall@gmail.com

Pradeep Kumar Singh
topksingh@gmail.com

[1] Madan Mohan Malaviya University of Technology, Gorakhpur, U.P., India

importance for challenging datasets. The issue of shallow vs. deep networks has long been argued for a long time in DLMs. Training of very deep learning models raises the issue of diminishing feature reuse (Zagoruyko and Komodakis 2017). This makes the training process cumbersome to train these models. González-Villà, S., Oliver et al. design a deep learning model with a perfect balance between resolution, depth, and width to achieve better accuracy and less error in training loss.

The public health sector is a highly critical sector where health professionals perform most interpretations of medical data. Advanced study of deep learning models reduced the complexity of the analysis of various medical images (Razzak et al. 2018). Image enhancement is an essential component of preprocessing. It is therefore important to examine the association between image improvement and the deep learning approach. He, K., Zhang, X., Ren et al. the authors employed on grayscale ImageNet dataset to pretrain, the Inception-V3 model tested on single-channel medical chest X-ray dataset outperformed both in terms of accuracy and speed. The DLMs ResNet-50 and DenseNet-161 approach a transfer learning methodology for the histopathology Kimia Path24 dataset, color, and grayscale image dataset. The DenseNet-161 uses a grayscale dataset, and the ResNet-50 uses a color dataset (Talo 2019). A novel DLMS implements split-transform-merge block (STM) and RE-based feature extraction to detect COVID-19 pneumonia (Shereen et al. 2020). Deep learning automated detection of medical imaging has shown promising results (Dou et al. 2021).

We propose a BoostNet DLM approach to improve the performance of musculoskeletal radiographs X-ray images. The highlight of this research is to assess the impact of three different image enhancement techniques (CLAHE, HEF, and UM) on DLMs for medical musculoskeletal radiographs X-ray images so that it helps to provide better opportunities for radiologists and orthopedic to ease out their lives with faster and more accurate results. The paper is organized into six different sections as follows: the most crucial related works described in the second section. In the third section, we have discussed the materials and methods used for the proposed model. In the fourth section, we have elaborated on the proposed model. In the fifth section, we have explained the simulation results and validation in detail. In the last section, we have discussed the conclusion and future scope.

2 Related work

Howard, A. G. et al. DLM is to investigate the efficiency of the model by increasing the depth (16–19 layers) on dataset ImageNet Challenge 2014. As the model layer extends to the 19th layer, the error rate of the deep learning model is saturated. To ease out the training of substantially deeper networks, the authors have developed a residual learning model (He et al. 2016a). The model has achieved an error rate of 3.57 % on the ImageNet test dataset (He et al. 2016b). In (He et al. 2016a) authors have suggested the propagation formulation for a deep learning model to transmit backward and forward pass directly from block to block. In Howard et al. (2017) two hyper-parameter resolution and width multiplier, the model creator can develop the best size model based on the problem, that is, a constraint. In (Tan and Le 2019) the authors suggested an appropriate scaling approach, which uniformly scales all three parameters width/depth/resolution dimensions using a compound coefficient (Jairath et al. 2021).

The authors have applied a transfer learning approach to both the deep learning models DenseNet-161 and ResNet-50 without a fully connected layer (Talo 2019). The research work was carried out on the Kimia Path24 dataset, both grayscale and colored format. The DenseNet-161 utilizes a grayscale dataset to achieve a classification accuracy of 97.89%, and the ResNet-50 utilizes a color dataset to achieve a classification accuracy of 98.87%. The authors Jaderberg, M. et al. Jaderberg, M. et al. presented a new version of the ResNet model. In this model, the authors have eliminated the global average pooling layer and added an adaptive drop-out. The Montgomery County Chest X-ray to achieve a classification accuracy of 87.71%, NIH X-ray set to achieve a classification accuracy of 62.9%, and the Shenzhen chest X-ray to attain a classification accuracy of 81.8%.

The authors Triwijoyo et al. (2020) have worked on the STARE dataset. The dataset is resized into three different datasets 31 × 35, 46 × 53, and 61 × 70 pixels classified with 15 different eye diseases. The studies have shown that input datasets with size 31 × 35 and 61 × 71 pixels have achieved the highest training accuracy and the input test dataset with size 31 × 35 with an accuracy of 80.93%. The authors Mahbod et al. (2020) study dermoscopic image dataset with different resolution sizes ranging from 64 × 64 to 768 × 768 pixels. Various deep learning models trained on DenseNet-121, ResNet-50, and ResNet-18. The author Dorffner, G., & Ellinger, I. et al. concludes the work as the classification performance significantly reduced on small-sized dataset 64x64 pixels and shows significant improvement with the dataset with size 128×128 pixels.

In Shin and Jung (2013) the edge area was improved by applying a high-frequency pass (emphasis) filter to the X-ray medical imaging field. To enhanced the edge and contrast of the X-ray image, the Gaussian high-pass filter is used with the optimized value offseta = 0.5 and cutoff frequencya = 0.05. In González-Villà et al. (2020) author has proposed a two-stage fusion approach (m-NLSS and

m-JLF) for improving brain segmentation performance. In Sahu et al. (2019) author has designed a model to remove the noise (Wiener Filter, Median Filter, Average Filter, Weighted Median Filter, Gaussian Filter) and enhance (CLAHE) the color fundus image. The Weighted Median filter combines with the CLAHE technique gives a 7.85% improvement in Peak Signal to Noise Ratio (PSNR). P. K., & Yadav, D. et al. author has discussed a new nonlinear UM enhancement technique (NLUM). NLUM can help boost the diagnosis and treatment by increasing fine details in mammograms.

3 Materials and methods for the proposed model

3.1 Bone X-ray image dataset

The musculoskeletal radiograph (MURA) is a collection of a total of 40561 image and bone X-ray images. The dataset contains 55.63% normal and 44.36 % abnormal X-ray images. This dataset was published by Rajpurkar et al. (2017) the most popular X-ray dataset (Rakhra et al. 2021). We have reorganized the MURA dataset into musculoskeletal radiograph bone classification (MURA-BC) for our experiments. The data set is organized into two folders (train and test), and each folder contains seven subfolders for each study, shoulder, elbow, humerus, finger, wrist, and hand. Only normal X-rays were extracted from the MURA dataset. The MURA-BC X-ray dataset details listed in Table 1.

3.2 Deep learning benchmark models

The key technical points about benchmark deep learning models are discussed below:

- *Efficientnet B0:* In Tan and Le (2019) author has proposed the efficientnet model based on the scaling theory for deep learning models. The three scaling factors taken into account are depth, width, and resolution. We have implemented this efficientnet (efficientnet: B0) baseline model for our research purposes.
- *MobileNet:* MobileNet is a lightweight and effective model (Howard et al. 2017). This model is designed to overcome challenges on the hardware level, such as limited memory, energy, and power. The model was designed for depth-wise separable convolution. These hyper-parameters help the model builder to select the appropriate DLM size for the framework depending on the problem constraints.
- *ResNet18:* In He et al. (2016a) author have submitted the ResNet model to the ImageNet Competition (ILSVRC) in 2015, and 152 layers 8 × deeper VGG nets were a winner in the Image NetChallenge. Two essential features implemented are dropout and batch normalization. At the network edge, the architecture also lacks fully connected layers.
- VGG-19: The VGG-19 model contains 19 trainable layers including convolutional and fully connected layers as well as max pooling, and dropout (Simonyan and Zisserman 2014). The DLMs classify 1000 different object categories (mice, keyboards, pencils, and various animals, etc.). As a result, the DLM has mastered the rich features of classification for a wide range of images (Setiawan et al. 2019). The network has an image scale of 224-by-224. The research shows that network depth is an essential component for improved performance (Zhang et al. 2015). The drawback of VGGNet is that the assessment is costly higher and requires much more memory to handle 19.6 billion FLOPs and approx. 143 million parameters.

Table 1 MURA-BC X-ray dataset detail

Study	Train set	Test set
Elbow	2925	92
Finger	3138	92
Forearm	1164	69
Hand	4059	101
Humerus	673	68
Shoulder	4211	68
Wrist	5765	140
Total	21,935	630
Complete dataset size:22,565		

3.3 Image enhancement techniques

The mathematical strength of different types of image enhancement techniques is discussed below:

- *CLAHE:* CLAHE image enhancement technique (Zuiderveld 1994), Input image ($I_{Orignal}$) is divided into non-overlapping ($R_{\text{contextual}}$) regions called sub image, titles, and blocks (Setiawan et al. 2019). The CLAHE method has primarily two key parameters: Clip limit (C_{limit}) and non-overlapping regions ($R_{\text{contextual}}$). These two parameters mainly control the enhanced image quality. N_{av} is average number of pixels in each gray level calculated as depicted in (1).

$$N_{av} = \frac{N_{crX} \times N_{crY}}{N_g} \tag{1}$$

where N_g = Gray levels number in the $R_{\text{contextual}}$, N_{crX} = Pixels number in the x dimensions of $R_{\text{contextual}}$, N_{crY} =

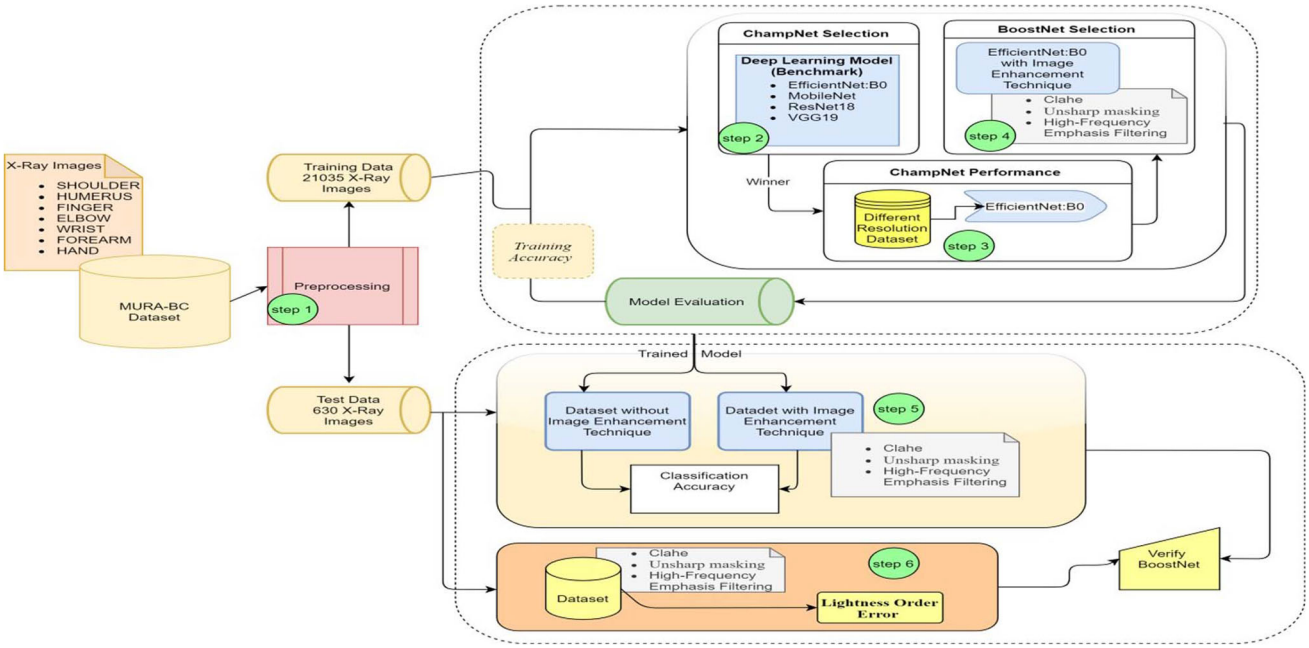

Fig. 1 Block design of our suggested paradigm

Pixels number in the y dimensions of $R_{\text{contextual}}$.the actual clip limit (C_{limit}) is computes as depicted in

$$C_{limit} = N_c \times N_{av} \tag{2}$$

where $N_c =$ maximum multiple of average pixels in gray level of $R_{\text{contextual}}$, $N\sum c$ = total number of clipped pixels. The number of pixels distributed averagely at each gray level is computed as depicted in (3).

$$N_{acis} = \frac{N\sum c}{N_g} \tag{3}$$

The Pd distributed pixel is computed as depicted in (4).

$$Pd = \frac{N_g}{N_{lp}} \tag{4}$$

where N_{lp} denotes the remaining number of clipped pixels.

- *HEF:* HEF is an enhancement technique that employs a Gaussian filter to enhance the edges in the input image (Bundy and Wallen 1984). The edges emerge presented in the high-frequency variety as they have more shifts that are dramatic in intensity (Deshmukh et al. 2021). This enhancement technique generates a low contrast-enhanced image and implements the Histogram Equalization method to improve contrast and sharpness. In the algorithm, the radius represents sharpness intensity. The original image is implemented through the Fourier transformation and the filter function. After the inverse transformation, we will have a filtered image. Secondly, the contrast of the image is in tune with histogram equalization. The Gaussian high pass filter is calculated as depicted in (5)

$$\text{Gau}\_{\text{filter}(x,y)} = 1 - e^{-D^2(x,y)/2D_0^2} \tag{5}$$

where D0 denotes the cut-off distance, and the F(i, j) denotes Fourier transform computed as depicted in (6)

$$F(i,j) = \sum_{x=0}^{h-1}\sum_{y=0}^{w-1} f(x,y)e^{-j2\pi\left(\frac{ix}{h}+\frac{jy}{w}\right)} \tag{6}$$

where x and i = 0, 1,2, ……h-1 and y and j = 0,1,2, ……w-1, F(x, y) denotes inverse Fourier transform computed as depict in (7):

$$F(x,y) = \frac{1}{hw}\sum_{x=0}^{h-1}\sum_{y=0}^{w-1} f(i,j)e^{-j2\pi\left(\frac{ix}{h}+\frac{jy}{w}\right)} \tag{7}$$

- *UM*: UM is an image enhancement technique that sharpens the original image (Polesel et al. 2000). The

Table 2 Enhancement techniques parameters details

Enhancement techniques	Parameters
CLAHE	Window Size:8 × 8
	Clip Limit:40
HEF	D0:70
UM	Radius:5
	Amount:2

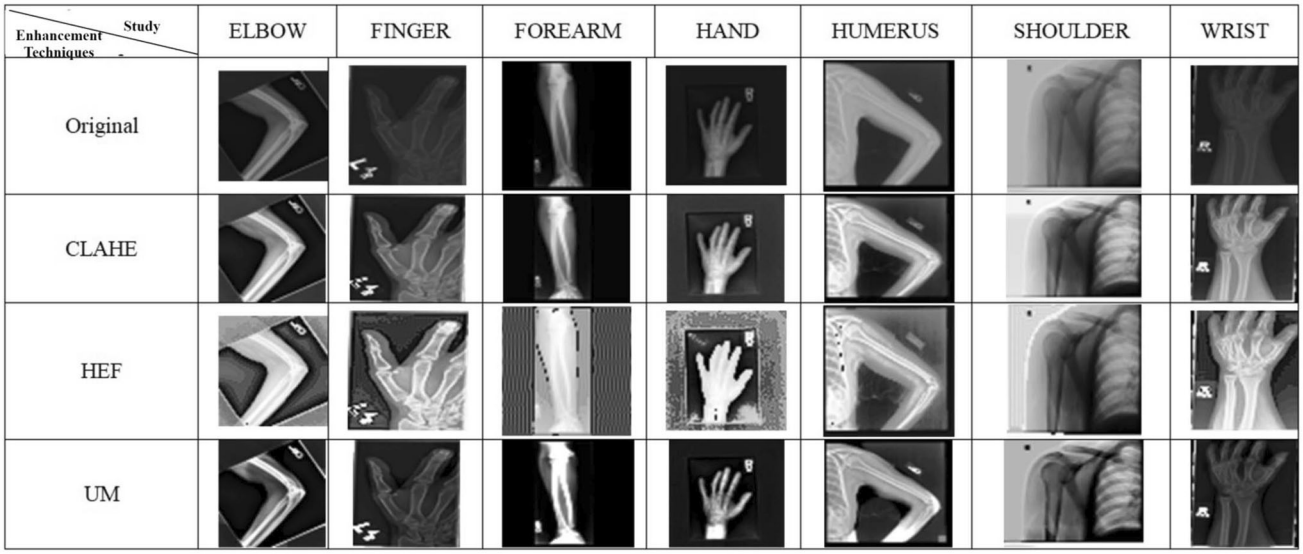

Fig. 2 Enhancement technique outcomes

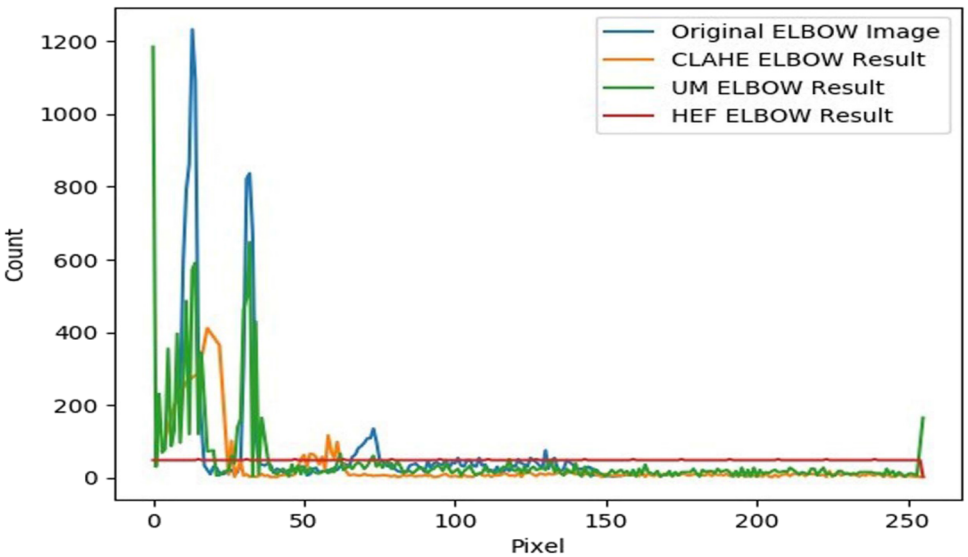

Fig. 3 Histogram of randomly Elbow image from MURA-BC

sharp details are calculated as the difference between the original and its Gaussian blur image. These collected details are then scaled and added back to the original image. At the beginning of this technique, Gaussian blur is applied to the input image (Shin and Jung 2013; Simonyan and Zisserman 2014). The radius and amount are two important parameters for Gaussian blur (Ramponi 1998). The size of the edge to be increased is affected by the radius. The amount is a factor of lightness or darkness contrast is added to the edges obtained through the equation as depicted in (8).

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{8}$$

where x and y denote the horizontal and vertical distance from the source, σ denotes the Gaussian distribution standard deviation. $I\_{enhanced}$ is an enhanced image obtained through the equation as depicted in (9)

$$I\_{enhanced} = I\_{orignal} + \text{contrast\_value} * (I\_{blur}) \tag{9}$$

where $I\_{orignal}$ the original image, and I_{blur} unsharp image.

Table 3 Training Accuracy of benchmark deep learning models

Training accuracy				
Epoch	EfficientNetB0	MobileNet	ResNet18	VGG19
1	59.0755	59.333955	44.50047	29.52292
2	78.02305	61.427714	54.92331	34.61147
3	83.78104	69.646071	67.78663	38.75943
4	85.91645	72.885423	78.62046	47.31802
5	87.42056	75.404919	81.35328	55.43174
6	88.42471	76.404919	83.64122	63.57088
7	89.38649	76.704919	86.36556	71.13804
8	89.91611	78.781459	87.30192	76.17575
9	90.15761	79.184163	88.89077	79.3026
10	90.31014	81.184163	89.39073	82.72604
11	90.48809	81.884163	89.69579	85.3148
12	90.6957	83.598848	90.61097	86.8062
13	90.98805	84.223155	90.89484	87.8273
14	90.95839	86.763834	91.31853	89.19583
15	91.30158	88.034912	91.80154	89.60681
16	91.67443	89.263622	92.01763	90.19998
17	91.84815	90.36946	92.01863	90.53046
18	91.65749	90.615202	92.02763	90.844
19	92.01339	91.640539	92.05999	91.96578
20	92.12355	91.424456	92.05999	91.96678
Max %	92.12355	91.640539	92.05999	91.96678

Table 4 Test Accuracy of benchmark deep learning models

Test accuracy				
Epoch	EfficientNetB0	MobileNet	ResNet18	VGG19
1	74.02519	68.32633	54.64907	23.93521
2	80.20396	78.16437	49.13017	33.23335
3	79.90402	82.96341	72.88542	24.35513
4	81.46371	83.62328	43.31134	43.97121
5	82.78344	86.92262	61.42771	50.56989
6	84.94301	87.70246	69.64607	55.72885
7	82.0036	88.84223	84.22316	59.56809
8	86.56269	88.42232	85.4829	50.08998
9	89.0222	89.92202	86.14277	70.60588
10	88.06239	90.76185	86.68266	72.70546
11	89.0222	90.64187	87.22256	65.08698
12	85.4829	91.18176	89.86203	67.30654
13	86.38272	87.40252	90.10198	78.22436
14	85.24295	86.38272	87.16257	82.60348
15	87.94241	90.04199	89.14217	86.14277
16	90.16197	90.10198	88.96221	88.36233
17	90.94181	89.68206	89.74205	88.06239
18	84.46311	91.60168	90.16197	89.80204
19	91.30174	90.40192	90.40192	87.76245
20	86.5027	91.78164	90.50178	88.18236
Max %	91.30174	91.78164	90.50178	89.80204

4 Proposed work

There are four major phases of the proposed model: Image preprocessing, benchmark DLMs training from scratch & validation, ChampNet processed with different resolution datasets and applied image enhancement techniques. The highlights of the proposed model are to select the ChampNet from the benchmark deep learning model and implement image enhancement techniques to boost the ChampNet performance. Lightness Order Error (LoE) validates the performance of BoostNet = ChampNet + image enhancement technique. Figure 1 depicts the block design of our suggested paradigm.

4.1 Research environment

The research work is carried out in the virtual environment. The host virtual machine is equipped with Ubuntu operating system, 12 GB RAM, and six virtual CPUs from the Intel Xeon silver 2.10 GHZ processor server. Python 3.0 is used for the implementation of the proposed model.

4.2 Image preprocessing

The Pre-processing X-ray images improve the key details of the raw image. The image-preprocessing two-step process dataset generation and transformation.

4.3 Dataset generation

In this phase, The MURA-BC X-ray dataset has been used in various pixel estimation dataset generation such as 32×32, 40×40, 48×48, 56×56, 64×64, 72×72, 80×80, 88×88 pixels. The MURA-BC X-ray dataset arranges in two folders: train and test. The train folder contains a total no of 21,935 and the test folder contains a total no of 630 X-ray images from seven different classes.

4.4 Transformation

In this phase, the training dataset is randomly cropped with 4 padding and X-ray images are randomly flipped horizontally. This technique provides an edge on the test dataset. The unseen X-ray image dataset captured for the test set can be in a random fashion. The normalization method is used to reduce unwanted noise or distortion

Table 5 Training error rate of benchmark deep learning models

Training error rate				
Epoch	EfficientNetB0	MobileNet	ResNet18	VGG19
1	1.383322	1.21302	1.731406	2.199559
2	0.642499	1.086369	1.353959	1.704481
3	0.476108	0.86186	0.914831	1.592747
4	0.418697	0.792313	0.628165	1.433213
5	0.380908	0.704619	0.545471	1.227937
6	0.352331	0.676819	0.487505	1.025349
7	0.333074	0.644619	0.40772	0.844052
8	0.315124	0.625471	0.387088	0.706743
9	0.311794	0.595194	0.339379	0.606142
10	0.303991	0.575274	0.329567	0.505616
11	0.300183	0.545334	0.311932	0.444331
12	0.292185	0.484095	0.292552	0.400369
13	0.287271	0.45264	0.280182	0.37328
14	0.2814	0.402656	0.276089	0.352496
15	0.275123	0.370232	0.25803	0.335794
16	0.262874	0.333399	0.249875	0.319002
17	0.260169	0.306982	0.249775	0.307171
18	0.259444	0.295539	0.249675	0.302845
19	0.241755	0.269972	0.242157	0.267149
20	0.244397	0.275724	0.242157	0.268149
Min %	0.241755	0.269972	0.242157	0.267149

Table 6 Test error rate of benchmark deep learning models

Test Error Rate				
Epoch	EfficientNetB0	MobileNet	ResNet18	VGG19
1	0.726903	0.866627	1.180009	2.032817
2	0.567598	0.694435	1.907381	1.749024
3	0.575895	0.505958	0.792313	1.926473
4	0.52121	0.478591	2.260284	1.456533
5	0.518982	0.39099	1.086369	1.378928
6	0.429103	0.381977	0.86186	1.092998
7	2.691613	0.352362	0.45264	1.096194
8	0.415552	0.353442	0.445867	1.425752
9	0.338509	0.333816	0.432549	0.868752
10	0.371915	0.303375	0.39823	0.784532
11	0.345888	0.319616	0.423801	1.163236
12	0.401998	0.291593	0.319726	1.09825
13	0.958694	0.388062	0.296771	0.652425
14	0.441828	0.411231	0.37014	0.554189
15	0.375713	0.317347	0.350389	0.452177
16	0.300061	0.296771	0.334783	0.377894
17	0.327415	0.33928	0.306369	0.40352
18	0.432823	0.326823	0.312	0.324585
19	0.276527	0.298766	0.298766	0.416029
20	0.413443	0.329408	0.293466	0.383946
Min %	0.276527	0.291593	0.293466	0.324585

signals. The X-ray image captured through the image modality system may be incomplete and devoid of essential details, such as irregular staining and poor contrast (Q et al. 2020).

4.5 Benchmark deep learning models training

The training of the benchmark deep learning model (EfficientNet: B0, MobileNet, ResNet18, VGG19) has been performed from scratch. The MURA-BC 32 × 32 X-ray image dataset has been used for training, validation, and testing purposes. This phase will help us to determine the best model from the benchmark is deep learning model (Mahajan et al. 2021).

4.6 ChampNet processed with different resolution datasets

In this phase, we have determined the performance of the deep learning model processed with different resolution datasets 40 × 40, 48 × 48, 56 × 56, 64 × 64, 72 × 72, 80 × 80, 88 × 88 pixels. The experiment performed in this phase is to select the resolution of the dataset for which the model performance gets stable in terms of training time and accuracy.

4.7 Image enhancement techniques

The Image enhancement phase is an essential aspect of our proposed model. The main aim of this phase is to boost the performance of the deep learning model and figure out the best enhancement techniques for the bone X-ray images. In this paper, total three enhancement techniques, CLAHE, HEF, UM are implemented. Table 2 contains the detailed parameters of the enhancement techniques. In Fig. 2, we demonstrate the outcomes of enhancement techniques implemented on some of the original X-ray images and Fig. 3 shows a histogram of randomly selected elbow images from the MURA-BC dataset.

5 Simulation, results, and validation

From the above applied methodologies, it can be seen that the histogram for randomly Elbow image from MURA-BC for different models shows some variation which can be detected by simulating the above results and then

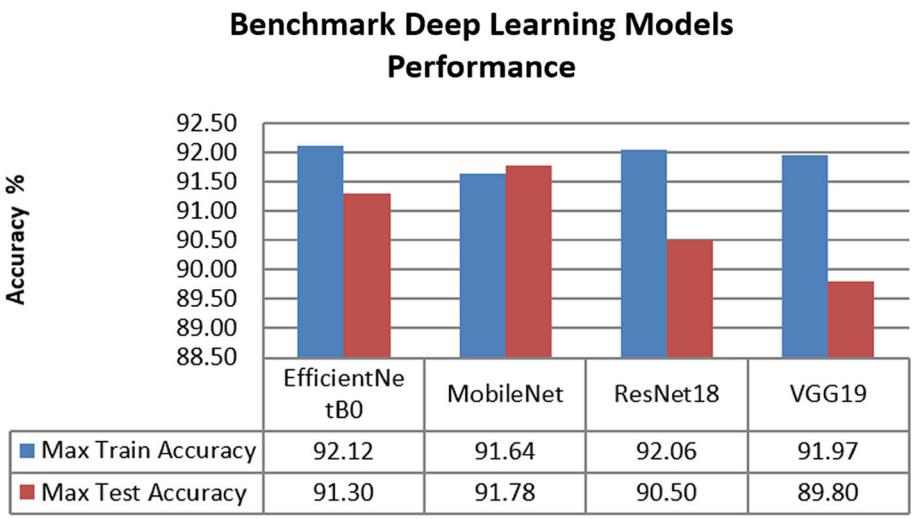

Fig. 4 Max train and test accuracy of different DLMs

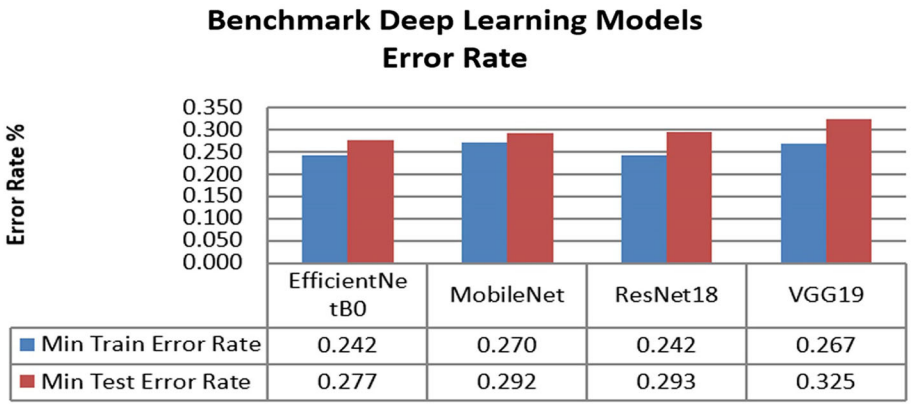

Fig. 5 Train and test error rate of different DLMs

comparing the results to find best one. The simulation of the proposed model is categorized in three major phases: benchmark deep learning model training & validation, ChampNet processed with different resolution dataset, implementation of image enhancement techniques on ChampNet, and the simulation of these three phases has been performed on Python 3.0. The model evaluation and validation explained in subsections.

5.1 Model evaluations

The performance of our model evaluated using accuracy (A_{cc}) and Cross-entropy error rate (E_r). Validation is performed through Lightness Order Error (LoE).

- **Accuracy**

The accuracy is the total number of correctly classified images out of the total number of the images in the dataset (Mall et al. 2019). The 'Accuracy' computes as depicted in (10) are as follows:

$$A_{cc} = (TP + TN)/(TotalNumber) \tag{10}$$

The 'total number' of the images in the dataset is computed as depicted in (11) as follows:

$$TotalNumber = TP + TN + FP + FN \tag{11}$$

where TP denotes True positives, TN denotes True negative, FP denotes False positive, False denotes negative FN.

- *Cross-entropy* Cross-entropy is widely used in the deep learning training process (Zhang and Sabuncu 2018). The Cross-entropy function is computed as depicted in (12):

$$E_r = (ylog(p) + (1-y)\log(1-p)) - (ylog(p) + (1-y)\log(1-p)) \tag{12}$$

Losses calculated separated for each class per observation and sum of the result computed as depicted in (13):

$$-\sum_{c=1}^{M} y_{obs,c} log(po, c) \tag{13}$$

where M > 2 denotes multiclass classification, log is the natural log, y is a binary indicator (0 or 1), c is the correct

Table 7 ChampNet accuracy based on different resolution train dataset

Training Report								
Epoch	EfficientNetB0 32 × 32	EfficientNetB0 40 × 40	EfficientNetB0 48 × 48	EfficientNetB0 56 × 56	EfficientNetB0 64 × 64	EfficientNetB0 72 × 72	EfficientNetB0 80 × 80	EfficientNetB0 88 × 88
1	59.07	61.08	53.57	46.95	58.43	60.63	68.70	69.13
2	78.02	83.02	79.70	76.30	80.55	81.85	87.81	88.27
3	83.78	87.41	86.66	83.37	87.36	88.31	90.78	91.28
4	85.91	89.31	88.59	86.32	89.61	90.72	92.44	92.45
5	87.42	90.08	89.66	88.15	90.86	92.18	93.09	93.39
6	88.42	90.77	90.68	89.11	91.36	92.72	93.55	93.30
7	89.38	91.21	91.19	89.84	91.82	91.26	93.65	94.10
8	89.91	91.48	91.25	90.25	92.42	92.91	93.90	93.91
9	90.15	91.76	91.73	90.64	92.22	93.19	93.93	94.24
10	90.31	92.05	91.69	91.00	92.85	93.40	93.85	94.18
11	90.48	91.79	91.85	91.39	92.47	93.48	94.22	94.44
12	90.69	92.45	91.91	91.56	93.00	93.89	94.39	94.31
13	90.98	92.63	91.84	91.69	92.35	93.50	94.39	94.45
14	90.95	92.63	92.29	91.75	92.61	93.77	94.08	94.17
15	91.30	92.62	91.53	92.01	92.86	93.86	94.22	94.25
16	91.67	93.07	91.98	92.18	93.03	93.19	94.19	94.47
17	91.84	92.50	92.33	92.14	92.83	93.69	94.53	94.38
18	91.65	92.65	91.99	92.30	92.89	93.71	94.55	94.86
19	92.01	92.94	92.54	92.30	92.94	93.64	94.64	94.50
20	92.12	92.74	92.33	92.55	93.10	93.70	94.39	94.2
Max %	92.12	93.07	92.54	92.55	93.10	93.89	94.64	94.86

Table 8 ChampNet accuracy based on different resolution test datasets

Testing report								
Epoch	EfficientNetB0 32 × 32	EfficientNetB0 40 × 40	EfficientNetB0 48 × 48	EfficientNetB0 56 × 56	EfficientNetB0 64 × 64	EfficientNetB0 72 × 72	EfficientNetB0 80 × 80	EfficientNetB0 88 × 88
1	74.02	61.08	68.86	67.84	71.02	72.76	80.44	80.44
2	80.20	83.02	79.72	67.30	81.82	81.40	86.20	86.74
3	79.90	87.41	81.94	78.76	85.84	87.94	88.78	90.40
4	81.46	89.31	86.92	79.60	89.26	88.72	90.70	92.38
5	82.78	90.08	87.52	85.96	89.44	91.12	89.92	87.22
6	84.94	90.77	89.20	82.48	89.56	87.82	91.24	90.46
7	82.00	91.21	89.68	84.76	86.26	89.02	91.24	91.18
8	86.56	91.48	88.66	87.40	87.22	91.84	91.66	92.86
9	89.02	91.76	87.16	87.70	89.80	90.04	91.78	91.42
10	88.06	92.05	89.74	88.90	87.88	92.02	91.30	90.76
11	89.02	91.79	87.88	89.02	90.58	91.24	89.92	91.60
12	85.48	92.45	84.16	83.80	89.68	90.70	91.54	91.54
13	86.38	92.63	85.36	88.36	89.02	91.78	89.92	92.32
14	85.24	92.63	89.08	90.46	88.96	91.90	91.96	91.84
15	87.94	92.62	86.98	89.14	90.94	88.00	89.44	91.00
16	90.16	93.07	88.36	90.76	88.54	87.16	94.24	88.48
17	90.94	92.50	86.92	86.86	91.30	91.42	92.68	90.82
18	84.46	92.65	85.90	89.92	91.96	88.06	89.20	90.10
19	91.30	92.94	87.40	89.38	88.60	91.36	92.80	92.14
20	86.50	92.74	87.70	89.92	85.90	91.4	93.46	90.82
Max %	91.30	93.07	89.74	90.76	91.96	92.02	94.24	92.86

Table 9 ChampNet training duration on different resolution datasets

ChampNet with different resolution	Training duration for 20 epochs
EfficientNetB0 32 × 32	30 Min
EfficientNetB0 40 × 40	55 Min
EfficientNetB0 48 × 48	1 h 50 Min
EfficientNetB0 56 × 56	2 h 45 Min
EfficientNetB0 64 × 64	3 h 20 Min
EfficientNetB0 72 × 72	5 h
EfficientNetB0 80 × 80	7 h 50 Min
EfficientNetB0 88 × 88	10 h

classification for observation, p denotes the predicted probability for observation of class c.

- *Lightness Order Error* The naturalness is crucial for image enhancement technique, but most of the techniques cannot maintain the naturalness effectively. We have considered the well-known image quality assessment (IQA) technique as Lightness Order Error (LoE) (Wang et al. 2013). This IQA technique provides the foremost solution among the methods (HEF, UM, and CLAHE) tested. LoE measure is based on the differences between the original input image (I_{input}) and enhanced image ($I\_{enhanced}$).the low LoE score indicates best solution and preserves the naturalness in enhanced images. The LoE is computed as depict in (14):

$$\mathrm{LoE} = \frac{1}{\mathrm{h} * \mathrm{w}} \sum_{\mathrm{i}=1}^{\mathrm{h}} \sum_{\mathrm{j}=1}^{\mathrm{w}} \mathrm{RD}(\mathrm{i},\mathrm{j}) \tag{14}$$

When h and w are the height and the width of RD, (x,y) is the relative order difference. In equation (15), the relative order difference is defined for the original image and the enhanced image.

$$\begin{aligned}\mathrm{RD}(\mathrm{I},\mathrm{J}) = &\sum_{\mathrm{i}=1}^{\mathrm{h}} \sum_{\mathrm{j}=1}^{\mathrm{w}} (\mathrm{U}(\mathrm{L}(\mathrm{x},\mathrm{y}),\mathrm{L}(\mathrm{i},\mathrm{j})) \\ &\oplus (\mathrm{U}(\mathrm{L}_{-\mathrm{enhance}}(\mathrm{x},\mathrm{y}),\mathrm{L}_{-\mathrm{enhance}}(\mathrm{i},\mathrm{j})))\end{aligned} \tag{15}$$

where L(x, y) lightness is computed as depicted in (16) and U (x, y), the unit step method computes as depicted in (17):

$$\mathrm{L}(\mathrm{x},\mathrm{y}) = \max_{\mathrm{c}\in[\mathrm{r},\mathrm{g},\mathrm{b}]} \mathrm{I}^{\mathrm{c}}(\mathrm{x},\mathrm{y}) \tag{16}$$

$$\mathrm{U}(\mathrm{x},\mathrm{y}) = \begin{cases} 1, x \geq y \\ 0, else \end{cases} \tag{17}$$

5.2 Experiment results for ChampNet selection

We have implemented our proposed model in Python, which provides the pathway to boost the performance of DLMs to classify bone X-ray images, we have performed a sequence of different experiments to analyze and confirm the effectiveness of our proposed model on the benchmark medical image dataset. To verify the efficiency of our model on MURA-BC medical imaging dataset, first, the training of the benchmark deep learning model (EfficientNet: B0, MobileNet, ResNet18, VGG19) has been performed from scratch (Rizwan et al. 2008). The MURA-BC 32 × 32 X-ray image dataset has been used for model training, validation, and testing purposes. The model training was performed for 20 epochs. Table 3 contains the results of training accuracy of benchmark deep learning model EfficientNet: B0, MobileNet, ResNet18, VGG19. We have obtained max training accuracy values of 92.12%, 91.64%, 92.05%, and 91.96%, respectively.

The test accuracy of EfficientNet: B0, MobileNet, Resnet18, and VGG19 benchmark models are depicted in Table 4. The max test accuracy achieved 92.12%, 91.64%, 92.05%, and 91.96%, respectively.

Table 5 contains the training error rate of benchmark deep learning model EfficientNet: B0, MobileNet, ResNet18, and VGG19. We obtain min training error rate values of 0.2814, 0.402656, 0.276089, and 0.3524966, respectively.

The test accuracy of EfficientNet: B0, MobileNet, Resnet18, and VGG19 benchmark models are depicted in Table 6. The min test error rate achieved 0.276527, 0.291593, 0.293466, and 0.324585, respectively.

The ChampNet selection is based on two standards: the max accuracy and the min error rate. The ChampNet is selected based on the training and test accuracy indicated in Fig. 4. Maximum training and test accuracy are 92.15 and 91.30, respectively.

With the help of Fig. 5, based on the training and testing error rates of various deep learning models. We have estimated the minimum training and testing error rates as 0.242 and 0.277, respectively. Through Figs. 4 and 5, we can easily determine EfficientNet: B0 as our ChampNet from the benchmark DLMs.

Table 10 ChampNet training and test accuracy for different enhancement techniques

	ChampNet + CLAHE	ChampNet + HEF	ChampNet + UM	ChampNet + CLAHE	ChampNet + HEF	ChampNet + UM	ChampNet + CLAHE	ChampNet + HEF	ChampNet + UM
	Training accuracy on train dataset			Test accuracy with enhancement on test dataset			Test accuracy without enhancement on test dataset		
1	75.35	62.30	66.64	68.74	67.13	79.75	75.35	18.42	65.15
2	91.08	79.62	85.91	86.03	73.05	81.97	73.01	18.96	70.19
3	92.93	85.43	89.53	85.19	80.18	89.57	81.64	18.42	76.66
4	93.62	88.37	91.09	87.29	85.39	88.68	83.08	19.68	67.13
5	94.07	89.47	92.14	90.53	80.06	85.32	87.76	20.58	65.15
6	93.99	90.51	92.83	88.31	85.03	89.87	88.66	19.98	72.77
7	94.30	91.04	93.54	89.60	77.90	91.67	85.72	18.90	77.20
8	94.07	91.67	93.42	90.59	88.08	88.98	88.84	21.06	68.45
9	94.28	91.59	93.85	82.43	84.73	87.42	86.56	34.55	76.72
10	94.43	91.98	93.59	89.12	85.03	91.85	87.34	35.75	72.95
11	94.31	93.01	93.96	91.07	78.44	90.71	90.34	22.74	70.55
12	94.38	92.91	93.94	92.42	81.08	83.28	88.78	35.93	75.52
13	94.61	93.16	94.23	91.55	79.58	89.04	90.70	40.13	65.33
14	94.60	93.32	94.42	92.54	81.08	93.53	86.80	23.22	72.35
15	94.79	93.21	94.20	92.93	82.34	90.35	91.06	25.73	65.51
16	94.70	94.74	95.26	91.85	92.93	94.61	92.02	28.19	63.47
17	94.87	94.99	95.27	94.99	91.20	92.45	87.28	24.06	75.46
18	94.83	95.27	94.98	92.45	90.90	94.39	94.18	28.49	84.58
19	95.01	95.06	95.00	92.03	93.77	93.41	90.22	24.12	69.59
20	95.88	95.57	95.69	91.55	94.79	93.11	91.90	20.34	69.71
Max %	95.88	95.57	95.69	94.99	94.79	94.61	94.18	40.13	84.58

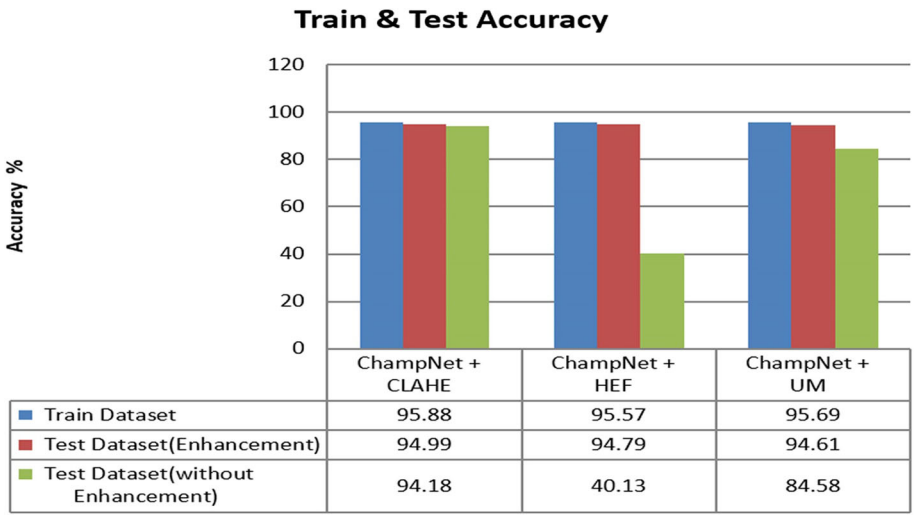

Fig. 6 Train and test accuracy on the train, test (with and without enhancement) dataset

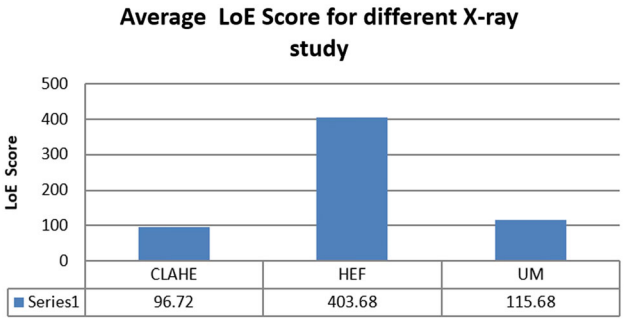

Fig. 7 LoE score for different X-ray studies

Table 11 LoE Scores for different X-ray study datasets

LoE Score	CLAHE	HEF	UM
LoE_ELBOW	116.19	407.99	138.88
LoE_FINGER	79.52	613.34	203.26
LoE_FOREARM	125.11	324.60	54.16
LoE_HAND	37.82	487.82	131.14
LoE_HUMERUS	89.25	585.77	104.15
LoE_SHOULDER	170.83	1.59	95.30
LoE_WRIST	58.36	404.67	82.84
Average_LoE Score	96.72	403.68	115.68

5.3 Experiment results for ChampNet processed with different resolution datasets

First, this experiment contributes to finding out the relationship between deep learning model performance and different dataset resolutions (Prasanalakshmi and Farouk 2019). Second, estimate the training time at different resolutions. The results in Tables 7 and 8 show the different resolution X-ray image dataset with the classification performance based on train and test accuracy, respectively. Table 9 contains estimate training time on different resolution datasets. However, the performance of DLMs improves with the growth in dataset resolution. As we increase the dataset resolution, the training time also increases. Thus, it is evident that from Tables 7, 8, and 9 that the 64 × 64-pixel resolution dataset is the best in terms of accuracy and training duration. Due to this reason, we have selected a 64 × 64-pixel X-ray image dataset, further investigating the performance of the model in the next phase.

5.4 Experiment results for ChampNet processed with different image enhancement

In this phase of research, the main objective is to boost the performance of DLMs. We have processed the dataset finalized from the previous phase with different image enhancement techniques, namely, CLAHE, HEF, and UM. The results in Table 10 and Fig. 6 show the classification performance of ChampNet with enhancement techniques on the train the test dataset with and without enhancement techniques applied (Awotunde et al. 2021; Munirathinam et al. 2021). All three techniques of image enhancement on the training dataset perform approximately the same in the range of 95.88%. The difference was examined during the test dataset with and without enhancement techniques (Zaman et al. 2018). From Table 10, it is clear that the CLAHE technique outperforms the other two techniques on both the test dataset with and without enhancement techniques. The CLAHE technique achieved 94.99% and 94.18% accuracy on the test dataset with and without enhancement techniques, respectively (Tang and Shabaz 2021). HEF achieved 94.79% and UM achieved 94.61% for the test dataset with enhancement techniques. HEF achieved only 40.13% and UM achieved 84.58% for the

test dataset without enhancement techniques (Sathya et al. 2019).

5.5 Experiment result validation with LoE

In the last phase, the LoE method is used to validate the result of the previous phase of the experiment as depicted in Fig. 7. The LoE score for different bone X-ray images is shown in Table 11. The low LoE score indicates the best solution and preserves the naturalness in enhanced images (Manzoor et al. 2018). The CLAHE LoE score is 96.72, which is the lowest among the other two techniques. The HEF LoE score is 403.68 and the UM LoE score is 115.68. The low LoE score of CLAHE validates the result of the previous phase.

6 Conclusion

The proposed design of BoostNet can enhance the performance of the MURA-BC dataset. The BoostNet plays a dynamic role in advancing the state-of-art performance of the musculoskeletal radiograph datasets. In this paper, we have introduced a model to improve DLMs in the medical imaging domain. Specifically, we have executed a series of experiments to enhance the accuracy of the model and authenticate the results using LoE techniques. The obtained findings are interesting for a wide variety of reasons:

- Benchmark DLMs (EfficientNet: B0, MobileNet, ResNet18, VGG19). The EfficientNet: B0 (ChampNet) outperforms other deep learning models, MobileNet, ResNet18, VGG19 in terms of high accuracy and low error rate. The EfficientNet: B0 provides high performance with minimum hardware resources. The experiment was performed in a virtual environment with 12 GB RAM and six virtual CPUs from an Intel Xeon Silver 2.10 GHZ processor server.
- The ChampNet performance improves gradually as we increase the dataset resolution, but the performance of the model gets stable. The EfficientNet: B0 with 64 × 64 resolution achieves stable accuracy (Tables7 and 8). This finding helps to improve the model accuracy.
- The ChampNet with 64 × 64 resolution dataset is implemented with three different enhancement techniques (CLAHE, HEF, and UM). The CLAHE outperforms the other two enhancement techniques, HEF and UM (Table 10). The ChampNet with CLAHE technique is referred to as BoostNet.
- The outcome of the ChampNet (64 × 64 resolution) with different enhancement techniques (CLAHE, HEF, and UM) experiment is verified with the LoE technique (Table 11 and Fig. 6).

The outcome of the research is musculoskeletal radiograph X-ray images processed with CLAHE enhancement technique with DLMs. The BoostNet can be implemented on several other medical imaging problems. In the future, the experiment could be processed with a higher resolution dataset and high-performance hardware resources. This model provides an immediate, complete tool to guide medical professionals in the treatment process in multiple medical domains. The practical deployments and application order to respond to resource constraints. In future it helps to provide better opportunities for radiologists and orthopedic to ease out their lives with faster and more accurate results.

Funding This research received no external funding.

Declarations

Conflict of interest The authors declare that they have no conflict of interest and all ethical issues including human or animal participation has been done. No such consent is applicable.

References

Awotunde JB, Chakraborty C, Adeniyi AE (2021) Intrusion detection in industrial internet of things network-based on deep learning model with rule-based feature selection. Wirel Commun Mobile Comput

Bundy A, Wallen L (1984) High-emphasis filtering. In: Catalogue of artificial intelligence tools. Springer, Berlin, Heidelberg. pp 47–47

Deshmukh S, Thirupathi Rao K, Shabaz M (2021) Collaborative learning based straggler prevention in large-scale distributed computing framework. In: Kaur M (ed) Security and communication networks. Hindawi Limited. Vol. 2021. pp 1–9. https://doi.org/10.1155/2021/8340925

Dou C, Zheng L, Wang W, Shabaz M (2021) Evaluation of urban environmental and economic coordination based on discrete mathematical model. In: Singh D (ed) Mathematical problems in engineering. Hindawi Limited. Vol. 2021. pp 1–11. https://doi.org/10.1155/2021/1566538

González-Villà S, Oliver A, Huo Y, Lladó X, Landman BA (2020) A fully automated pipeline for brain structure segmentation in multiple sclerosis. NeuroImage: Clin 27:102306

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 770–778

He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. In: European conference on computer vision. Springer, Cham. pp. 630–645

Howard AG et al (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications, arXiv preprint http://arxiv.org/abs/1704.04861

Jaderberg M et al (2015) Spatial transformer networks, arXiv preprint http://arxiv.org/abs/1704.04861

Jairath K, Singh N, Jagota V, Shabaz M (2021) Compact ultrawide band metamaterial-inspired split ring resonator structure loaded band notched antenna. In: Kumar V (ed) Mathematical problems in engineering. Hindawi Limited. Vol. 2021. pp 1–12. https://doi.org/10.1155/2021/5174455

Krizhevsky A, Hinton G (2009) Learning multiple layers of features from tiny images

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25:1097–1105

Mahajan K, Garg U, Shabaz M (2021) CPIDM: a clustering-based profound iterating deep learning model for HSI segmentation. In: Shanmuganathan V (ed) Wireless communications and mobile computing. Hindawi Limited. Vol. 2021. pp 1–12. https://doi.org/10.1155/2021/7279260

Mahbod A, Schaefer G, Wang C, Ecker R, Dorffner G, Ellinger I (2021) Investigating and exploiting image resolution for transfer learning-based skin lesion classification. In: 2020 25th international conference on pattern recognition (ICPR). IEEE. pp 4047–4053

Mall PK, Singh PK, Yadav D (2019) Glcm based feature extraction and medical x-ray image classification using machine learning techniques. In: 2019 IEEE conference on information and communication technology. IEEE. pp 1–6

Manzoor U, Rizwan A, Demirbas A, Hafiz NAS (2018) Analysis of perception gap between employers and fresh engineering graduates about employability skills: a case study of Pakistan. Int J Eng Educ 34(1):248–255

Munirathinam R, Ponnan S, Chakraborty C, & Umathurai S (2021) Improved performance on seizure detection in an automated electroencephalogram signal under evolution by extracting entropy feature. Multimed Tools Appl, 1–16

Panetta K, Zhou Y, Agaian S, Jia H (2011) Nonlinear unsharp masking for mammogram enhancement. IEEE Trans Inf Technol Biomed 15(6):918–928

Polesel A, Ramponi G, Mathews VJ (2000) Image enhancement via adaptive unsharp masking. IEEE Trans Image Process 9(3):505–510

Prasanalakshmi B, Farouk A (2019) Classification and prediction of student academic performance in king Khalid University-A machine learning approach. Indian J Sci Technol 12:14

Q et al (2020) Zhan A GPU-based residual network for medical image classification in smart medicine. Inf Sci 536:91–100

Rajpurkar P, Irvin J, Bagul A, Ding D, Duan T, Mehta H, Ng AY (2017) Mura: Large dataset for abnormality detection in musculoskeletal radiographs. arXiv preprint arXiv:1712.06957

Rakhra M, Singh R, Lohani TK, Shabaz M (2021) Metaheuristic and machine learning-based smart engine for renting and sharing of agriculture equipment. In: Singh D (ed) Mathematical problems in engineering. Hindawi Limited. Vol. 2021. pp 1–13. https://doi.org/10.1155/2021/5561065

Ramponi G (1998) A cubic unsharp masking technique for contrast enhancement. Signal Process 67(2):211–222

Razzak MI, Naz S, Zaib A (2018) Deep learning for medical image processing: overview, challenges and the future. Classification in BioApps, pp 323–350

Rizwan A, Alvi MS, Hammouda MM (2008) Analysis of factors affecting the satisfaction levels of engineering students. Int J Eng Edu 24(4):811–816

Sahu S et al (2019) An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. Opt Laser Technol 110:87–98

Sathya D, Ganesh Kumar P, Prasanalakshmi B (2019) Enhancement of data security with reduced energy consumption in WMSN

Setiawan F, Yahya BN, Lee S-L (2019) Deep activity recognition on imaging sensor data. Electron Lett 55(17):928–931

Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R (2020) COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. J Adv Res 24:91

Shin CH, Jung CY (2013) An enhancement of medical image using optimized high-frequency emphasis filter. J Korea Inst Inf Commun Eng 17(3):698–704

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition, arXiv preprint http://arxiv.org/abs/1704.04861

Talo M (2019) Automated classification of histopathology images using transfer learning. Artif Intell Med 101:101743

Tan M, Le Q (2019) Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning. pp 6105–6114. PMLR

Tang S, Shabaz M (2021) A new face image recognition algorithm based on cerebellum-basal ganglia mechanism. In: Chakraborty C (ed) Journal of healthcare engineering. Hindawi Limited. Vol. 2021. pp 1–11. https://doi.org/10.1155/2021/3688881

Triwijoyo BK, Sabarguna BS, Budiharto W, Abdurachman E (2020) Deep learning approach for classification of eye diseases based on color fundus images. In: Diabetes and fundus OCT. Elsevier. pp 25–57

Wang S, Zheng J, Hu HM, Li B (2013) Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Trans Image Process 22(9):3538–3548

Zaman S, Chakraborty C, Mehajabin N, Mamun-Or-Rashid M, Razzaque MA (2018) A deep learning based device authentication scheme using channel state information. In: 2018 International conference on innovation in engineering and technology (ICIET). IEEE. pp 1–5

Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. Springer, Cham, pp 818–833

Zhang X et al (2015) Accelerating very deep convolutional networks for classification and detection. IEEE Trans Pattern Anal Mach Intell 38(10):1943–1955

Zhang Z Sabuncu MR (2018) Generalized cross entropy loss for training deep neural networks with noisy labels'. arXiv preprint http://arxiv.org/abs/1704.04861

Zuiderveld K (1994) Contrast limited adaptive histogram equalization. Graphics gems, pp 474–485