Journal of Engineering Sciences and Information Technology
Volume (6), Issue (7) : 30 Dec 2022
P: 21 - 34
ISSN: 2522-3321
مجلة العلوم الهندسية وتكنولوجيا المعلومات
المجلد (6)، العدد (7) : 30 ديسمبر 2022م
ص: 21 - 34

# Optimal deep learning network for musculoskeletal X-Ray imaging classification for hospital medical record system

**Naif Olayan Alresheedi**
**Mohammed Omair Alresheedi**
**Mohammed Ali Ahmed Alodhaybi**
College of Applied Medical Sciences || Majmaah University || KSA
Maternity & Children Hospital in Buraydah || Ministry of Health || KSA

**Abstract:** Background: Medical image archiving is one of the integral components of any hospital medical record system (HMRS). It includes, but is not limited to, MRI, CT-Scan, X-ray, Ultrasound, Musculoskeletal X-rays etc. The musculoskeletal X-ray images are relatively significant in number among the other types of medical imaging. Most of the existing HMRS use either the manual annotation of the images or use metadata for every image for archiving. This approach is found to be deficient because of intensive manual work, chances of misclassification, and reliance on human expertise. Moreover, archiving the images and their metafiles is relatively difficult to handle.

Methodology: This issue can be handled by a hybrid solution of computer vision and deep learning. In the recent literature, researchers have proposed using machine learning and deep learning algorithms for biomedical image classification and archiving. However, the literature is found to be insufficient to recommend a unified deep learning network for Musculoskeletal X-ray Image classification with greater accuracy and efficiency. The LERA dataset is considered one of the benchmark Musculoskeletal X-rays image datasets.

Results: To the best of knowledge the investigation of the best candidate of deep neural network is still missing in the literature. This study will present the logical and empirical rationale for the recommendation of the optimal deep learning network for X-ray Imaging Classification for Hospital Medical Record System using LERA (musculoskeletal radiographs) dataset. It has been concluded that the variants of Resnet, Google Net, and DarkNet are the suggested candidates for LERA x-ray image classification.

**Keywords:** Deep learning, Musculoskeletal X-ray Imaging Classification, hospital management system, model optimization, LERA dataset, machine learning.

# شبكة التعليم العميق المثلى لتصنيف التصوير بالأشعة السينية للعضلات الهيكلية لنظام السجلات الطبية بالمستشفى

**نايف عليان الرشيدي**
**محمد بن عمير الرشيدي**
**محمد بن علي العضيبي**
كلية العلوم الطبية التطبيقية || جامعة المجمعة || المملكة العربية السعودية

مستشفى الولادة والأطفال ببريدة || وزارة الصحة || المملكة العربية السعودية

**المستخلص:** أرشفة الصور الطبية هي أحد المكونات الأساسية لأي نظام للسجلات الطبية بالمستشفى (HMRS). وهي تشمل ، على سبيل المثال ، التصوير بالرنين المغناطيسي ، والأشعة المقطعية ، والأشعة السينية ، والموجات فوق الصوتية ، والأشعة السينية للعضلات الهيكلية وما إلى ذلك. تعد صور الأشعة السينية للعضلات الهيكلية ذات أهمية نسبية من حيث العدد بين الأنواع الأخرى من التصوير الطبي. تستخدم معظم نظام السجلات الطبية الموجودة بالمستشفى إما بالتعليق التوضيحي اليدوي للصور أو تستخدم البيانات الوصفية لكل صورة للأرشفة. وجد هذا النهج على انه ناقصاً بسبب العمل اليدوي المكثف ، وفرص سوء التصنيف ، والاعتماد التام على الخبرة البشرية. علاوة على ذلك ، من الصعب نسبيًا معالجة أرشفة الصور وملفات التعريف الخاصة بها.

المنهجية : يمكن معالجة هذه المشكلة من خلال الحل المختلط لرؤية الكمبيوتر والتعلم العميق. في الأدبيات الحديثة ، اقترح الباحثون استخدام خوارزميات التعلم الآلي وخوارزميات التعلم العميق لتصنيف الصور الطبية الحيوية وأرشفتها. ومع ذلك ، وجد أن الأدبيات ناقصة للتوصية بشبكة تعلم عميق موحدة لتصنيف صور السينية للعضلات الهيكلية بدقة وكفاءة أكبر. تعتبر مجموعة بيانات LERA واحدة من مجموعة بيانات صور الأشعة السينية للعضلات الهيكلية المعيارية.

النتائج على حد علمي ، لا يزال البحث عن أفضل مرشح للشبكة العصبية العميقة مفقودا في الأدبيات ، وستقدم هذه الدراسة الأساس المنطقي والتجريبي للتوصية بشبكة التعلم العميق المثلى لتصنيف التصوير بالأشعة السينية لنظام السجلات الطبية بالمستشفى باستخدام مجموعة بيانات LERA (الصور الإشعاعية للعضلات الهيكلية). تم الاستنتاج أن المتغيرات من Resent و Google Net و Dark Netهي المرشح المقترح لتصنيف صور الأشعة السينية LEAR .

**الكلمات المفتاحية:** التعلم العميق ، تصنيف التصوير بالأشعة السينية للعضلات الهيكلية ، نظام إدارة المستشفى ، تحسين النموذج ، مجموعة بيانات LERA ، التعلم الآلي.

## 1.1 Introduction :

During the last decade, there has been a rapid increase in the use of medical imaging. Researchers and scientists have made major contributions to improving the diagnostic and treatment flexibility and accuracy of medical imaging [1]. The scenario appears to be beneficial, yet it has resulted in the massive number of images set. This has necessitated the development of a reliable system for archiving these enormous image sets [2]. Most of the available solutions are using either the manual annotation of image labels or using metadata of image files [3]. In the recent literature, researchers have proposed to use machine learning algorithms and deep learning algorithms for biomedical image classification and archiving [4]. However, the literature is insuffiecient to recommend a unified deep learning network for Musculoskeletal X-ray Image classification with greater accuracy and efficiency [5]. This study will submit a comprehensive performance evaluation of approx. 19 deep learning networks on the resent benchmark dataset of the Musculoskeletal X-ray Image of LERA Dataset or MURA Dataset.

In particular, the selection and hyper-parameter optimization of the best candidate of a deep learning algorithm for musculoskeletal x-ray image classification will also be covered in the scope of this study [6]. The performance evaluation will be a function of training accuracy, loss function, and testing accuracy. Matlab 2021 will be used as a simulation tool with an integrated deep network designer app and experiment manager. The dataset will be divided into 70% training and 30% testing dataset. The best candidate of the deep learning network will classify the musculoskeletal x-ray image such as elbows, fingers, forearms, hands, humerus, shoulder, and wrist. The proposed deep learning network will be out-performed in terms of efficiency and efficacy among other peer deep learning networks for the given dataset of musculoskeletal x-ray image classification. Moreover, it would be able to classify the musculoskeletal x-ray image in a real-time manner. Finally, the hyper-parameter optimization of the best candidate will further increase its rank.

### 1.2 Problem statement

A comprehensive review of the radiography analysis and classification literature helps to bring the research closer to the following research question: "How to detect and classify the Musculoskeletal X-ray Imaging for Hospital Medical Record System with a higher degree of accuracy and tolerance" .

### 1.3 Hypothesis

OPTIMAL DEEP LEARNING NETWORK HAS GREAT INFLUENCE FOR FOR MUSCULOSKELETAL X-RAY IMAGING CLASSIFICATION FOR HOSPITAL MEDICAL RECORD SYSTEM.

### 1.4 Research Objectives

The research issue can be addressed through measurable research goals:

1- Establish an empirical parametric evaluation of a broad set of deep learning algorithms on the LERA data set for a relatively higher degree of accuracy and tolerance.
2- Recommendation of the best candidate of the deep learning algorithm on LERA for the ultimate degree of precision and tolerance.

### 1.5 Research importance

This scope of this study can further extend towards designing of a novel deep learning model for the higher degree of accuracy and minimum loss. Moreover, this model can also be deployed on the intelligent embedded systems like Resberry pi or coral AI for real time classification. This study was constaints as it need an intensive amount of computation cost during the training. In addition, a massive data is indeed a pressing need.

### 1.6 Fundamentals of Deep Learninng

This section will cover the principles of deep learning parameters. Each pre-trained network's performance is evaluated as a function of accuracy, loss, and training time in this study. In addition, each pre-trained network has a collection of input layers, convolutional layers, pooling layers, soft-max layers, activation function, fully connected layers, optimizer, and other components. This section contains a brief description of each parameter.

**Accuracy:** It is the ratio between true positive scores to the total number of sampling units.

**Activation function:** It is an activatoin function to the artificial neural network in order to define the decision function. Some of the activation functions are following Adam, Adagrad, Adamax, Adadelta, Nadam, RMSprop, Nesterov, Rprop - advanced gradient descent algorithms

**Average pooling:** This pooling method submits the average of the felted mraktd values

**AUC – ROC :** This curve is a performance measurement for classification challenges at numerous threshold settings. *ROC* is a probability curve and *AUC* depicts a level or measure of selectivity. It tells the efficiency of a model to differentiate among classes.

**Batch size:** It reflects the number of training instance, in each epoch.

**Back-propagation:** It is the learning method of the neural network where the error from the output layer is propagating backward to the hidden layers to update the weight. Primarily, the desired value of the hidden layer is not defined, therefore the updated weight and the error gradient of the output layer help~~s~~ to update the weight and error radient of the hidden layer.

**CNN:** Convolutional network - It is a classic deep learning network. 2D or 3D images are fed as input to the network, then weights and biases are assigned to various features of image. Afterwards the scores are in order to classify it.

**Confusion matrix:** It is the tabular illustration of the performance parameters like true possible, false positive and false negative. The diagonal value represents the true positive.

**Epoch:** It is referred to as the training iteration in the deep learning. One epoch means a complete dataset is processed in forwarding and backward directions through the neural network only once.

**Layer:** A layer is the top-level key-element in deep learning. A layer is a sack that ordinarily fed by the weighted input, applied non-linear activation functions and then delivers the generated output to the next layer

**Loss function**: It calculates the difference between the results of the neural network and the expected output.

**Softmax:** This function generates a vector as output that depicts the probability scores of a list of possible results.

### 1.7 Thesis Organization

This dissertation is organized in a logical manner. Chapter 1 presents an introduction of the domain of X-ray image classification as well as the use of a deep learning algorithm to classify x-ray images. Chapter 2 presents the literature review of the proposed research work along with the identification of the research gap analysis. Chapter 3 presents the proposed methodology and dataset description. The simulation results and analyses are depicted in Chapter 4. Finally, in Chapter 5, the conclusion work is presented.

### 1.1 Literature Review

It has been observed that intensive work on medical imaging using deep learning has been reported in the literature. It's not only the research and exploration, but also for the clinical diagnosis at the application layer [7]. In this section, a comprehensive literature review will be discussed along with the scope of work and limitations. Finally, this section will summarise the open areas in connection with the research gap from the literature.

Jaiswa et.al [8], proposed a deep learning network based on RCNN to diagnose pneumonia. This network has been trained on the Chest Radiograph reference dataset. One author referred to it as Mask-RCNN, derived from ResNET101 and ResNET50. The author noted that there had been significant improvements in the performance of the proposed model. Another author in [9] opted for a related approach in which CNN-based AlexNet and GoogleNET were used to form the system on the National Library of Medicine (PubMed) database. In the same domain, other CNN variants such as Inception- V3, self-control, BiLSTM, and FCNN have been developed for the diagnosis of ischemic or non-ischemic cardiomyopathy [10].

Deep learning variants have also been studied with a view to classifying malignant and benign cells [11]. In this work, the author used the pre-formed network of GoogleNet based on CNN, VCCNet, and ResNet for classifying cells. Another author in [12] suggested the use of ResNet 101 for the classification of skin lesions. More precisely, the author supports the diagnosis of melanoma dermoscopy. Likewise, other research has proposed using AlexNet, VGG16, and ResNet-18 for the classification of skin lesions [12].

In addition to clinical diagnosis, variants of the new deep network concept have been proposed for classifying and archiving biomedical images. In that context, a synergistic approach based on deep learning (SDL) has been proposed in the literature. This model has received training in pathology and modality data sets [11]. As well, Google Inception V3 is also reported in the literature for the histopathologic classification of breast cancer images [13]. In recent literature, DenseNet has also been nominated for the classification of biomedical images. Similar to [13], 3D DenseNet was formed on two data sets. The first is the data set collected in Haudong Hospital, and the other is the TCIA data set for pulmonary adenocarcinoma with EGFR mutation. Benign paroxysmal positional vertigo (BPPV) testing was also performed using CNN [44].

In [15] a further variation of CNN, i.e. a separate deep CNN was proposed for the screening of oral cancer from HSI images. Another researcher in [12] offered to use AlexNet and VGG-16 for the classification of medical sub-figures. This study alleges that Google Net and ResNet demonstrate greater accuracy than AlexNEt and VCC-16.

The use of deep learning has been observed to gain significant attention in the literature. In this context, many challenges have been proposed by the community of researchers. These challenges are primarily intended to address the biomedical problem within deep learning.

These challenges includes but are not limited to, Calcaneus Radiograph Investigation Technique, Diagnosis of finger joint destruction, Evaluation and diagnosis of hip fractures, Rheumatoid arthritis, detection, Categorization and recognition of calcaneus fracture, Musculoskeletal anomaly finding system, Diagnosing broken elbows in infants, A a computerized approach for marking fractured areas from infants musculoskeletal images, Human upper extremity radiographs identification and classification, and Cloud computation for anomaly finding in elbow radiograph images.

In the following sections, the proposed methodology for each of the challenges (mentioned above) will be reviewed. In addition, the open room in each contribution will also be highlighted.

In Calcaneus Radiograph Investigation [16], the presentation was presented to retrieve the characteristics using the SIFT descriptor and the subsequent classification was performed through CNN. This contribution indicated that robust variation of the overall approach may be the appropriate candidate for radiographic classification. In another challenge of diagnostic finger joint destruction [17], the cascading classifier methodology is chosen. That methodology is supplemented by CNN. This study found that the hybrid methodology of the waterfall classifier and CNN is one of the most suitable combinations. In [18], difficulties in diagnosing hip fractures have been initiated. The best proposal for this is from DCNN and DensNet.

This study also found that CNN variants with deep layers could be the best networks, but the complexity of the networks could be compromised. In the same year, the focus group introduced a challenge to detect rheumatoid arthritis. In this work, transfer learning and localization with the domain-specific data set were reported as the right streamlined solution [19]. In the area of calcaneal fracture categorization and recognition, ResNet and VGG are among the most appropriate variants of CNN [20]. Moreover, in the musculoskeletal abnormality the dilated DenseNet is found to be the optimum varcents of the deep neural network [21].

The deep learning approach is also advocated to be the right contribution to the analysis of biomedical images of infants and toddlers. For example, diagnose an elbow fracture in infants [22] and a fracture in the muscular skeletal images of the infant [23]. Again, in these fields, DCNN, Xception, and ResNet 18 has performed exceptionally well compared to other deep learning network variances.

In another reported work in the literature compliment the use of Multilayer CNN with integrated Hough transformation is found in thbestod combination for Human upper extremity radiographs identification [24], and ousnomaly finding in elbow radiograph [25].

The application of deep learning on the radiograph was primarily done in early 2000s when one of the researchers devised a natural language processing (NLP) based model for knowledge extraction related to Bacterial Pheumonia. Initially, the results were very significant as he submitted the 86% accuracy with 94% recall [26] as the transfer of learning and localization was also in practice with the pre-formed model. One of the researchers opted for the Chest Xpert dataset to classify control illness [28]. These studies were proposed by the University of Oxford and reported an accuracy of 30%.

In addition to ChestXpert, various other datasets were also included in the investigation, such as the acoustic scene TUT 2016 CheckXRay-14[34], MIMIC-CXR [27], LERA, and MURA. Researchers reported varying levels of accuracy and recall using methods such as CNN and LSTM variants [26] on these datasets. In the literature, a scholar suggested a Chest X Net. This is a CNN 121 layers which was formed on ChestXray 14 dataset. This dataset includes more than 100,000 coronal chest x-rays and 14 distinct diseases. In the same trend, another research was presented in [28] where the CNN was trained on two datasets, i.e., MNIST and CIFAR-10.

### 1.2 Gap Analysis

The comprehensive literature review provided an understanding of the importance of deep learning in analyzing and classifying X-rays. It was noted that the veracity of the Deep Learning Network has been explored across different datasets. Furthermore, learning to transfer to the pre-formed network was also studied using local data. Apparently, a great deal of work has been done on the analysis and classification of X-rays. However, very little work has been done on the recent LERA and MURA reference datasets. The research community has strongly voted and advocated for the LERA and MURA datasets for muskolosteleton radiography analysis. This is mainly due to the sheer volume of data and the validation of the dataset by the large group of experts in the field of knowledge. This created an urgent need to break the exploration of deep learning about the LERA and MURA data set with the variety of the pre-trained network. As a result, a stronger and better pre-trained network is expected. Furthermore, the hyper-parameter optimization of the proposed pre-formed network is also urgently needed and will increase the precision and reliability of the network based on deep learning. In this work, a comprehensive empirical study of many pre-Trained networks is proposed in this work. It involves submitting the selection for the best deep learning network. In addition, to refine and adapt the model, hyper-parameter

## 3. METHODOLOGY AND SYSTEM SETUP

### 3.1 Methodology

The scope of this study is bi-folded, i.e., experimental data acquisition and model design from the Optimum Deep Learning Network for Musculoskeletal X-ray Imaging Classification for Hospital Medical Record System. The experimental data was taken from the LERA-Lower Extremity RAdiographs reference data collection at Stanford University. Model design and evaluation, on the other hand, will be done with Matlab 2021 as a simulation tool.

The massive data set contains data from 182 patients who underwent x-ray examinations at Stanford University Medical Center between 2003 and 2014. The dataset is a collection of foot, knee, ankle, or hip images associated with each patient. The data instances will be divided into 70% training and 30% testing datasets. For the hospital's medical record system, the best candidate in the Deep Learning Network will classify the patient's musculoskeletal X-ray image.

In this study, a comprehensive parametric evaluation of a wider set of deep learning networks on the classification of musculoskeletal X-ray imaging from the hospital medical record system will be presented. The assessment will be based on the actual positive rate, false negative rate, confounding matrix, error coefficient, accuracy of training and accuracy of testing. Finally, the hyperparameter optimization of the best candidate will also include the scope of the study in order to continue its ascension. Figure 1 shows the proposed methodology for musculoskeletal X-ray classification using deep learning.

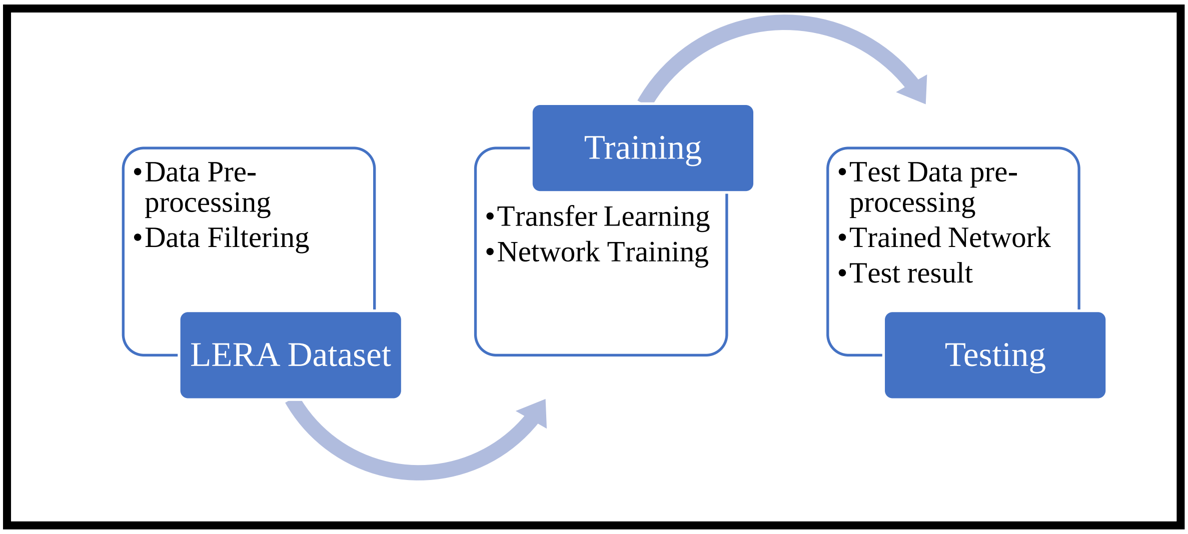


**Figure 1. Proposed Methodology Training and Testing**

### 3.2 Study Design and Dataset

LERA (Lower Extremity Radiographs) comprises a variety of articulations, bones and soft tissues. Musculoskeletal disorders (MSDs) of the ankle, knee, hip and foot. This data set was collected by HIPAA-complaint which compiled data from 182 patients. Between 2003 and 2014, these individuals were radiographically assessed at Stanford University Medical Centre.

This dataset includes a wide variety of images as it was compiled over a twelve-year period. Since the detection of musculoskeletal abnormalities is explicitly critical, and about 1.7 billion people around the world suffer from musculoskeletal disorders [56]. Therefore, it is vital to develop any computer-assisted tool with the help of deep learning techniques to provide a precise and correct diagnosis of a musculoskeletal disorder of multiple body parts in order to deliver the treatment on time can be delivered.

Figures 2 and 3 present the sample of images and the classification between the number of images for each class of the LERA data set.

**Figure 2. LERA Dataset sample of ANKLE, FOOT, HIP, and KNEE [6]**

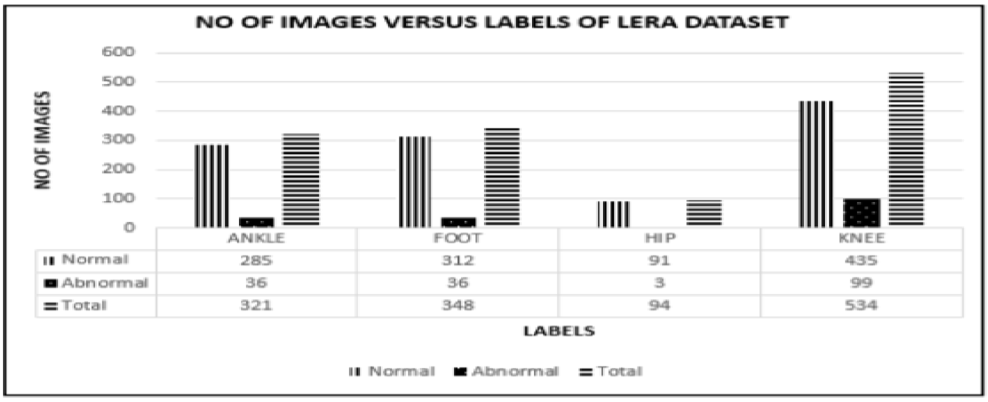


**Figure 3. Image Distribution of LERA Dataset**

### 3.3 Research outcomes

This study will contribute to a unified machine deep learning network that will outperform in terms of efficiency and efficacy among other peer deep learning network for the LERA dataset of patients' Lower Extremity Radiographs classification of the hospital record management system. Furthermore, the hyperparameter optimization of the best candidate will also encompass the scope of the study in order to continue its ascension.

### 3.4 Pre-trained deep Learning Network

This section will present the network layer architecture of pre-trained network. Moreover, the connection to the source and destination layer is also mentioned. In each subsection a separate table is presented to illustrate the network layer in this study.

## 3.5 SIMULATION RESULTS AND ANALYSIS

This Chapter will present a compereansive investigation and analysis of the simulation results. In this study a large number of pre-trained deep learning network were used. It includes, squeezenet, googlenet, inceptionv3, densenet201, mobilenetv2, resnet18, resnet50, resnet101, xception, inceptionresnetv2, shufflenet, nasnetmobile, nasnetlarge, darknet19, darknet53, efficientnetb0, alexnet, vgg16, and vgg19. The generally reported GPU utilization vs. accuracy is shown in Figure 6. Relatively, the same response has been observed in this study. The simulation was conducted in Matlab 2021 using Deep Network Designer App. The highest performance server machine was used for this simulation with the following system configurations:

| | |
|---|---|
| Processor | Intel® Xeon® CPU E5-2673 v3@2.4Ghz |
| RAM, | 8GB |
| OS | Windows Server |
| Graphics Card | NVIDIA GEFORCE RTX 2080 |

The dataset was first loaded into the server with the given distribution, as shown in Figure 5. It is to be noted that the HIP images are relatively unbalanced in the dataset, therefore, a slight data augmentation has been employed to eliminate the class imbalance issue.

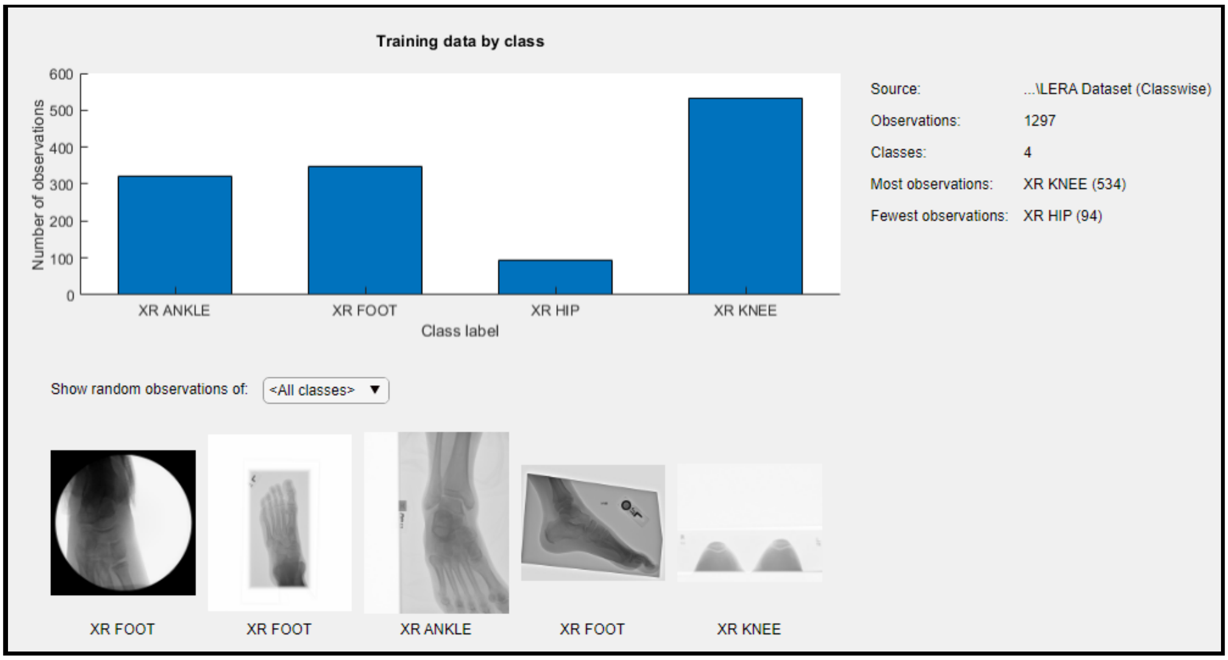


**Figure 5. Class wise distribution of Dataset**

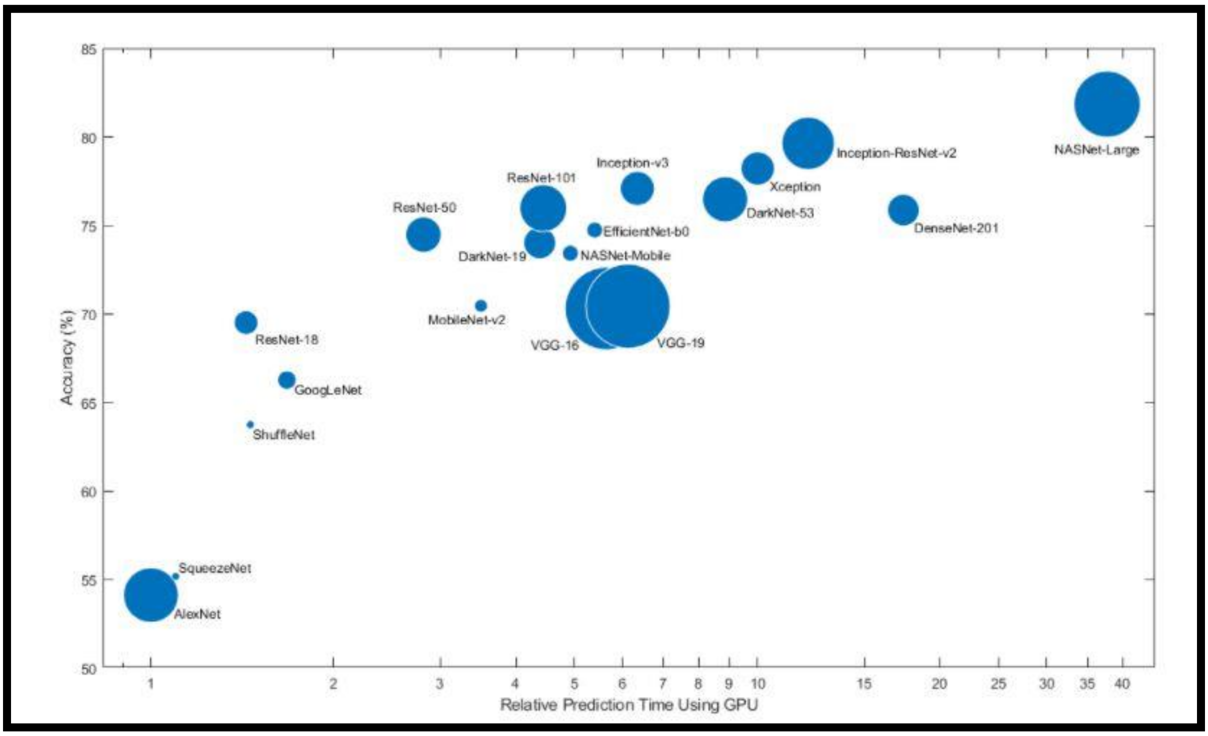


**Figure 6. GPU utilization vs. accuracy of pretrained network**

## 4. Results

### 4.1 Simulation Results

In this study, a wide variety of deep learning pretrained models have been investigated for the LERA dataset. The investigation is the function of the accuracy curve (blue curve), loss function orange curve. Moreover, the training time, epoch, learning rate and number of iterations for each network as shown in the respective figures. The desired outcome for the accuracy is the exponential rise in the accuracy curve with respect to the number of epochs and get plated new to the 100% of the accuracy. On the other hand, the exponential decay in the loss function is desired that attains the minimum value of loss. The Elapsed time in minutes, shows the total time required for training.

Figure 6 represents the SqueezNet with a very low accuracy of 25% with a training time of 13 min. This curve shows that the network has stopped training after the 10th iteration. Figure 7 is the training and loss curve of Google Net that attains acceptable accuracy in its 20th iteration. However, the total training time is 23 min. GoogleNet could be considered as one of the candidate for LERA image classification. Resonate 50 should have a relatively better response in Figure 8, where it attains the maximum accuracy in just the 10th iteration. Subsequently, the DarkNet 53 is far better as compared to ResNet 50 as it attains the highest degree of accuracy in just 5 iterations.

The response of DarkNet 19 is now highly acceptable as a candidate due to its unstable response and lower accuracy, as shown in Figure 10. Likewise, Figure 11 shows a similar response to Google net. The ShuffleNet attains the desired accuracy in its 10 iterations.

The NasNet Mobile and NasNet Large are structurally very complex, as is also evident from their layer layout and computational time which is 936 min. However, no substantial improvement can be observed in the performance. As indicated in Figures 14 and 15, both the Xception and Places365 Google appear to be the average performance.

The Mobile Net V2, DenseNet201, and Resnet 18 also have a similar response time as Darknet with an acceptable elapsed time. It has been concluded that the variants of Resnet, Google Net and DarkNet are the suggested candidates for LERA x-ray image classification.

### 4.2 Rationale of Parametric Evaluation

This section presents a comprehensive parametric evaluation of the LERA dataset X-ray classification using a large set of deep learning networks. The parameters include accuracy, and hearing loss. The comprehensive investigation of deep learning networks for x-ray classification has established a strong advocacy in favour of the most suited deep learning network for real-time classification for the said application.

## CONCLUSION

Medical imaging is now the essential element of the medical record system. A massive number of medical imaging reports like X-rays, CT-scans, ultrasounds, etc. are found in the medical record system. Due to the massive archive of these records, a robust content based image retrieval is a pressing need in the literature. Deep learning is reported as the best tool to handle large data with a higher degree of precision and accuracy. However, the network architecture may vary from application to application.

This study has submitted a unified deep learning network that will out-performed in terms of efficiency and efficacy among other peer deep learning networks for the LERA dataset of patients' Lower Extremity Radiographs classification in the hospital record management system. Furthermore, the hyper parameter optimization of the best candidates will also continue its ascension. It has been concluded that the variants of Resnet, Google Net and DarkNet are the suggested candidate for LEAR x-ray image classification.

## REFERENCES

[1] Tanzi, Leonardo, et al. "Hierarchical fracture classification of proximal femur X-ray images using a multistage deep learning approach." European Journal of Radiology 133 (2020): 109373.

[2] Chen, Xu, et al. "Automatic feature extraction in X-ray image based on deep learning approach for determination of bone age." Future Generation Computer Systems 110 (2020): 795-801.

[3] Tanzi, Leonardo, et al. "X-ray bone fracture classification using deep learning: a baseline for designing a reliable approach." Applied Sciences 10.4 (2020): 1507.

[4] Bhandary, Abhir, et al. "Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images." Pattern Recognition Letters 129 (2020): 271-278.

[5] Khalid, Hafsa, et al. "A comparative systematic literature review on knee bone reports from mri, x-rays and ct scans using deep learning and machine learning methodologies." Diagnostics 10.8 (2020): 518.

[6] LERA- Lower Extremity RAdiographs Standford University official website

[7] D. Shen, G. Wu, and H.-I. Suck, "Deep Learning in Medical Image Analysis," Annul Rev. Biomed. Eng., vol. 19, no. 1, pp. 221–248, 2017, doi: 10.1146/annurevbioeng-071516-044442.

[8] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, "Identifying pneumonia in chest X-rays: A deep learning approach," Meas. J. Int. Meas. Confed., vol. 145, pp. 511–518, 2019, doi: 10.1016/j. measurement. 2019. 05. 076.

[9] A. Fourcade and R. H. Khonsari, "Deep learning in medical image analysis: A third eye for doctors," J. Stomatol. Oral Maxillofac. Surg., vol. 120, no. 4, pp. 279–288, 2019, doi: 10.1016/j Jae rmas. 201 9.06.002.

[10] S. U. Khan, N. Islam, Z. Jan, I. Ud Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," Pattern Recognit. Lett., vol. 125, pp. 1–6, 2019, doi: 10.1016/j. Patike c.201 9.0 3.022. Page 133 of 153

[11] C. Xue, Q. Dou, X. Shi, H. Chen, and P. A. Heng, "Robust learning at noisy labeled medical images: Applied to skin lesion classification," Proc. - Int. Symp. Biomed. Imaging, vol. 2019-April, pp. 1280–1283, 2019, doi: 10.1109/ISBI.2019.8759203.

[41] J. Zhang, Y. Xie, Q. Wu, and Y. Xia, "Medical image classification using synergies deep learning," Med. Image Anal., vol. 54, pp. 10–19, 2019, doi: 10.1016/j. media. 2019. 02. 010.

[12] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinger, "Institute fr Pathophysiology and Allergy Researh , Medical University of Viena , Austria Department of Research and Developent , TissueGnosticsGmbH , Austria Department of Computer Sience , Loughborough Unierity , U . K . Department of Biomedical," pp. 1229–1233, 2019.

[13] W. Zhao et al., "Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning," Cancer Med., vol. 8, no. 7, pp. 3532–3543, 2019, doi: 10.1002/cam4. 2233.

[14] E.-C. Lim et al., "Developing a Diagnostic Decision Support System for Benign Paroxysmal Positional Vertigo Using a Deep-Learning Model," J. Clin. Med., vol. 8, no. 5, p. 633, 2019, doi: 10.3390/jcm8050633.

[15] P. R. Jeyaraj and E. R. Samuel Nadar, “Computer-assisted medical image classification for early diagnosis of oral cancer, employing deep learning algorithm,” J. Cancer Res. Clin. Oncol., vol. 145, no. 4, pp. 829–837, 2019, doi: Page 134 of 15310.1007/s00432-018-02834-7.

[16] J. Guo, W. Wang, H. Yan, J. Chen, H. Xu, and H. Li, “Calcaneus Radiograph Analysis System: Rotation-Invariant Landmarks Detectio , Calcaneal Angles Measurement and Fracture Identification.”

[17] T. Hirano et al., “Development and validation of a deep-learning model for scoring of radiographic finger joint destruction in rheumatoid arthritis.,” Rheumatol. Adv. Pract., vol. 3, no. 2, p. rkz047, 2019, doi: 10.1093/rap/rkz047.

[18] C. T. Cheng et al., “Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs,” Eur. Radiol., vol. 29, no. 10, pp. 5469–5477, 2019, doi: 10.1007/s00330-019-06167-y.

[19] J. Rohrbach, T. Reinhard, B. Sick, and O. Dürr, “Bone erosion scoring for rheumatoid arthritis with deep convolutional neural networks,” Comput. Electr. Eng., vol. 78, pp. 472–481, 2019, doi: 10.1016/j. compelli ng.20 19. 08.003.

[20] Y. D. Pranata et al., “Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images,” Comput. Methods Programs Biomed., vol. 171, pp. 27–37, 2019, doi: 10.1016/j. comb .2019 .02 .006.

[21] S. Li, “Deep adversarial model for musculoskeletal quality evaluation,” Inf. Process. Manag., vol. 57, no. 1, p. 102146, 2020, doi: 10.1016/j. ipm. 2019. 102146. Page 135 of 153

[22] J. C. Rayan, N. Reddy, J. H. Kahn, W. Zhang, and A. Annapragada, “Binomial Classification of Pediatric Elbow Fractures Using a Deep Learning Multiview Approach Emulating Radiologist Decision Making,” Radiol. Artif. Intell., vol. 1, no. 1, p. e180015, 2019, doi: 10.1148/ryai. 2019180015.

[23] P. H. Yi et al., “Automated semantic labeling of pediatric musculoskeletal radiographs using deep learning,” Pediatr. Radiol., vol. 49, no. 8, pp. 1066–1070, 2019, doi: 10.1007/s00247-019-04408-2.

[24] N. Pradhan, V. Singh Dhaka, and H. Chaudhary, “Classification of Human Bones Using Deep Convolutional Neural Network,” IOP Conf. Ser. Matter. SCI. Eng., vol. 594, no. 1, 2019, doi: 10.1088/1757-899X/594/1/012024.

[25] A. Galal, F. Hisham, M. Mohamed, S. Hassan, T. Ghanim, and A. Nabil, “Automatic recognition of Elbow musculoskeletal disorders using cloud application,” ACM Int. Conf. Proceeding Ser., pp. 80–84, 2019, doi: 10.1145/3328833.3328870.

[26] R. S. Evans, “Automatic Detection of Acute Bacterial Pneumonia from Chest X-ray Reports Automatic Detection of Acute Bacterial Pneumonia from Chest X-ray Reports,” no. November, 2000.

[27] W. Zhang, F. Zhang, Y. Zhang, and L. Qin, Cohesive subgraphs with hierarchical Page 161 of 185 decomposition of big graphs, vol. 11448 LNCS, no. 2016. Springer International Publishing, 2019.

[28] P. Rajpurkar et al., “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,” pp. 3–9, 2017.