Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEEP LEARNING-XCEPTION ALGORITHM FOR UPPER BONE ABNORMALITIES CLASSIFICATION

ALAA M. A. BARHOOM[1], MOHAMMED RASHEED J. AL-HIEALY[2], SAMY S. ABU-NASER[3]

[1]University Malaysia of Computer Science & Engineering (UNIMY), Cyberjaya, Malaysia

p05210001@student.unimy.edu.my

[2]University Malaysia of Computer Science & Engineering (UNIMY), Cyberjaya, Malaysia

[3]Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine

ABSTRACT

Upper bones are strong and flexible tissue made up of collagen and calcium phosphate. They mainly contribute to the movement of the human body and serve as a protective shield for the body's soft organs such as the brain, lungs, and the heart. Without these bones, the human body would not be constructed to function ordinarily. However occasionally, due to accidents, an individual is exposed to some diseases such as injury or infection that lead to defects in the regular shape and growth of bone construction. This deficiency in the bone structure is so-called bone abnormalities. Frequently, the preliminary diagnosis of bone abnormalities is made by specialists using X-rays of the patient's injury site to show the shape and density of the bones. They are classified into normal or abnormal. The detection and classification of bones depend on the experience and human effort. So the error in the results of this process can expose the patient to a great danger and catastrophe of his life. Therefore, deep learning algorithms from artificial intelligence were applied to help specialists avoid wrong or inaccurate diagnoses when detecting bone abnormalities in X-ray images by using a pre-trained convolutional neural network called Xception model. The model was customized to fit the bone abnormalities classification then applied to a dataset consisting of 42000 X-rays of the upper bones of some patients collected from Kaggle depository. We trained, validated, and tested the customized Xception model. The proposed Xception model attained Precision (85.20%), Recall (85.13%) and F1-Score (85.07%).

Keywords:- *Bone Abnormalities, Deep Learning, Xception*

1. INTRODUCTION

Bones are one of the most important parts of the human body that help the person move and perform various life functions. It also exists as the first line of defense for the soft parts of the human body, such as the rib cage, skull, and others. Bones are the densest natural structures in terms of composition, but any change in the composition of their structure and shape tends to render human functions incapable of performing in their natural form. Therefore, this change is called bone abnormalities. The main causes of bone abnormalities are genetic factors, direct injury, and infection of parts of the bone and muscle structure [1]-[3].

Bone fractures are considered to be one of the most common bone abnormalities in the world as recent studies have shown that in the United States of America more than half a million cases are hospitalized annually due to bone fractures, resulting in nearly 180,000 patients with fractures referred to aged care centers that are due to the low mass and density of the bones. At one age, the most dangerous of these injuries are hip fractures, which often result in death if not treated well. Therefore, bone abnormalities have a significant impact on people's physical life in terms of pain, deformity, and disability in performing vital and psychological functions, and neglecting their diagnosis and treatment at the right time is disastrous. Orthopedic malposition also has a negative effect on the psyche of the patient's relatives and can burden them financially [4]-[7].

Accurate and timely diagnosis of bone abnormalities in their early stages avoids many negative consequences for the patient, so doctors rely on bone X-rays made by radiological centers for diagnosis, evaluation, and classification because they are inexpensive and quick to obtain an X-ray image ofthe bone is preserved by directing a small amount of ionizing radiation onto a body part such as the hand, foot, thigh, elbow, wrist and other parts of the bony body using a machine. The bones absorb a lot of radiation because of their density, and the soft organs let the radiation through. The radiation efficiencies are recorded on photographic film so that the bones appear a color that is almost white and the soft organs. With an almost gray color, through this film, it is possible to look for defects and changes in the bone structure in order to detect bone abnormalities. X-ray images are often stored in the database of the health center or hospital so that the doctor can diagnose and treat various cases of illness [8]-[11].

It is easy to get an x-ray of the bones, but it is difficult to get an accurate and correct diagnosis in time to save the patient's life. The current problem in health centers is the reading of the X-rays by laymen and inexperienced medical professionals and the result is an incorrect diagnosis, also the problem of a long time it takes the specialist doctor to prepare the medical report, which reaches up to at least 20 minutes, in addition to the problems of X-ray image blur, from which it can be difficult to extract features, these problems threaten the patient in terms of safety, so artificial intelligence and machine learning sciences have been used to help specialists and patients alike [12]-[16].

In order to improve the quality of diagnosis and classification of X-ray images in the medical field, researchers were interested in applying and exploiting deep learning techniques for this purpose, many of them competed to propose models and methods to achieve advanced diagnosis and provide classification as results through feature extraction and classification by modeling contain multiple layers to data processing, then training on the dataset and testing, the greatest success in this is due to the convolutional neural network (CNN), which was proposed in the late 1990s, which achieved great popularity, especially in classifying images and extracting features from them with high accuracy and other two-dimensional data compared to other models [17]-[20].

In this study, the is to examine the efficiency of the convolutional neural network in diagnosing and classifying bone deformities in radiographs, the Xception model was chosen to apply to a dataset consisting of 42,000 samples for an X-ray of the upper bone obtained from Kaggle composed of 7 types of human bones (wrist, finger, elbow, hand, forearm, humerus, shoulder). Each type of bone either: normal and abnormal label; thus we have 14 different categories. Xception was configured for retraining, after processing the dataset using the Python programming language, the Keras library and the Tensor Flow platform in the Google Colab environment with GPU to get the results, the first is the bone type and then the classification of the type in terms of normal and abnormal [21].

Xception was chosen because it is a convolutional neural network that is pre-trained on a massive dataset of 14,000,000 images, and was awarded the highest award in 2015, containing 71-layer convolutional layers and weights that can detect and extract features from an image. Through the results, it is possible to predict in the future the possibility of suggesting a better model or not, so in this paper, we review the methodology and its application and record its results to compare it in the future with the performance of other models [22]-[25].

2. PROBLEM STATEMENT

Radiographs, most often called X-rays, produce shadow-like images of bones and certain organs and tissues. X-rays are very good at finding bone problems. They can show some organs and soft tissues, but MRI and CT scans often give better pictures of them. Still, X-rays are fast, easy to get, and cost less than other scans, so they might be used to get information quickly.

X-ray radiography detects bone fractures, certain tumors and other abnormal masses, pneumonia, some types of injuries, calcifications, foreign objects, dental problems, etc. When used appropriately, X-ray scans can diagnose possibly life-threatening conditions such as blocked blood vessels, bone cancer, and infections, and so we might save human lives.

Every day many people around the world suffer bone fractures, causing severe disabilities for the patient in the event that they are not diagnosed well, accurately and quickly before complications arise, so when deep learning is used to detect bone deformities and fractures, the result is: Reducing the burden on diagnostic centers X-rays and reduce the burden on doctors' staff to solve the problem of errors in the normal diagnosis, which is likely to have a greater error rate that may lead to disasters on the patient.

So, in this study, we hope to train a model using the dataset to help physicians for making diagnose of human bones in X-rays to detect bone abnormalities. The proposed model will use deep learning techniques to increase accuracy and efficiency in the diagnosis. These include human bones X-ray images, image processing techniques and data analysis.

3. OBJECTIVES

3.1 Main objective:

Implementation a software model used to detect bone abnormalities and classify some of Human bone if found in x-ray images by the following types:

- Bone types: 1- Humerus, 2- Forearm, 3- Shoulder, 4- Hand, 5- Finger, 6- Wrist, 7- elbow.
- Diagnostic result: 1- normal, 2- abnormality.

3.2. Specific objectives:

- Rapid diagnosis and detection of Human bone abnormalities.
- Reduce the cost of diagnosis and repetitive images.
- Increase proficiency using deep learning techniques to detect Human bone abnormalities.
- Reducing the burden on X-ray diagnostic centers, to also reduce the burden on medical staff in hospitals and health centers to be able to provide the best services to patients.

4. LIMITATION

The dataset collected from Kaggle depository for bone abnormalities are limited to the following bones: Humerus, Forearm, Shoulder, Hand, Finger, Wrist, and elbow.

The total number of labels in the dataset is 14. Each type of bone in the dataset is either abnormal or normal.

5. REVIEW OF LITERATURE

Many researchers have used the sciences of artificial intelligence and machine learning to support the medical field, as data of various types, text, and images, have been processed to create reports that help improve the efficiency and quality of medical work regarding diagnosis, prediction, and classification of some diseases that can affect people, and the main focus has been on training models (deep learning) on various patient datasets to achieve accurate results that can benefit decision-makers in health centers and hospitals. The authors in [19-20, 23, 25-39] used different pre-trained model to detect bone abnormalities as binary problem. That means they detect if there is an abnormality in the image or not.

Esteva et al. [19] works in "Object classification, localization, and detection, refer to identifying the type of an object in an image, the location of objects present, and both type and location simultaneously." The first working core with GPU powered deep learning approach, in 2012, by ImageNet Large-Scale Visual Recognition Challenge, the challenge, which included many researchers aspiring to improve and develop computer vision, this resulted to implementing deep learning models, all of which were successful and accurate results, especially in the medical field and disease detection, the accuracy of classification and detection of these models was similar to the results of the diagnosis of the ordinary doctor, and sometimes exceeded their levels, the greatest credit is due to the Convolutional Neural Network (CNN), which has the ability to extract differences and analyse data from a huge set of data that divided tasks and was able to collect similar images and other tasks, (CNN) type of deep learning algorithm which hardcodes translational invariance, a key feature of image data. But the challenges of the medical field need more work with and development of different models of deep learning.

Cernazanu et al. [20] were able to train a convolutional neural network on a dataset consisting of medical images such as X-ray images and CT images containing specific diseases, using some operations such as convolutional, fully connected layers, and pooling, the convolutional neural network receives the images as input, then it will convert images to flattened vectors in the end the softmax layer represent the "elements of the output vector, which actually represents the probability of detecting the disease in the images, during the training process, the internal parameters of the network layers are iteratively adjusted to improve accuracy. Typically, lower layers learn simple image features edges and basic shapes which influence the high-level representations," so that the training outputs are the answer to the question: is there a disease or not? For example, normal or abnormal.

Sitaula et al. [21] in light of the spreading epidemic of Corona virus disease, a group of researchers retrained a "deep learning model based on a convolutional neural network called VGG-16 by using the attention module, that can capture the spatial relationship between the ROIs in CXR images, which could identify the likely regions of COVID-19's effect in the human lungs, with appropriate convolution layer 4th pooling layer, they designed a novel deep learning model to perform fine-tuning in the classification process," three sets of images were used to make the evaluation process, and the results were satisfactory, and we can rely on deep learning models to detect infection with the Corona virus within lung images. The researchers relied on VGG16 model for two important reasons: "firstly, it extracts the features at low-level by using its smaller kernel size, secondly, it has a better feature extraction ability for the classification of COVID-19 CXR images, used the pre-trained weight of ImageNet. It helps to overcome the over-fitting problem as they had limited amount of COVID-19 CXR images for training purpose, used the four main building blocks in model: Attention module, Convolution module, FC-layers, and Softmax classifier," after training the model on the dataset with a total of 445 images divided into three training and testing categories, the accuracy of the model work was satisfactory by 79.58%. This indicates success VGG-16 Model Learning and training on a small and accurate set of data.

Esteva et al. [22], have "examine the strength of deep learning approaches for pathology detection in chest radiographs to explore the ability of CNN learned from a non-medical dataset to identify different types of pathologies in chest X-rays, after tested algorithm on a 433-image dataset, the best performance was achieved using CNN and GIST features, they got an accuracy rate 0.87-0.94 for the different pathologies. The results demonstrate the feasibility of detecting pathology in chest X-rays using deep learning approaches based on non-medical learning."

Gulshan et al. [11] they developed the "EyePACS-1 data set consisted of 9963 images from 4997 patients (mean age, 54.4 years; 62.2% women; prevalence of RDR, 683/8878 fully gradable images [7.8%]); the Messidor-2 data set had 1748 images from 874 patients (mean age, 57.6 years; 42.6% women; prevalence of RDR, 254/1745 fully gradable images [14.6%]). For detecting RDR, the algorithm had an area under the receiver operating curve of 0.991 (95% CI, 0.988-0.993) for EyePACS-1 and 0.990 (95% CI, 0.986-0.995) for Messidor-2. Using the first operating cut point with high specificity, for EyePACS-1, the sensitivity was 90.3% (95% CI, 87.5%-92.7%) and the specificity was 98.1% (95% CI, 97.8%-98.5%). For Messidor-2, the sensitivity was 87.0% (95% CI, 81.1%-91.0%) and the specificity was 98.5% (95% CI, 97.7%-99.1%). Using a second operating point with high sensitivity in the development set, for EyePACS-1 the sensitivity was 97.5% and specificity was 93.4% and for Messidor-2 the sensitivity was 96.1% and specificity was 93.9%."

Moreno et al. [15] "presented first findings towards assessing how computer vision, natural language processing and other systems could be correctly embedded in the clinicians' pathway to better aid on the fracture detection task. We present some initial experimental results using publicly available fracture datasets along with a handful of data provided by the National Healthcare System from the United Kingdom in a research initiative call. Results show that there is a high likelihood of applying transfer learning from different existing and pre-trained models (VGG16, Resnet50, InceptionV3) to the new records provided in the challenge, and that there are various ways in which these techniques can be embedded along the clinicians' pathway." The accuracy rate By VGG16 model was 92.7 % in one of the stages.

Moran et al. [25] introduced model to "classify regions in periapical examinations according to the presence of periodontal bone destruction. This study considered 1079 interproximal regions extracted from 467 periapical radiographs. This data was annotated by experts and used to train a ResNet and an Inception model, which were after evaluated with a test set. Inception presented the best results and an impressive rate of correctness even on the small and unbalanced dataset. The final accuracy, precision, recall, specificity, and negative predictive values are 0.817, 0.762, 0.923, 0.711, and 0.902, respectively. The ROC and PR curves also demonstrate the good performance of both models. These results suggest that the evaluated CNN model can be used as a clinical decision support tool to diagnose periodontal bone destruction in periapical exams."

El-Saadawy et al. [24] presents a method for detecting the "fractures in the seven extremity upper bones (shoulder, humerus, forearm, elbow, wrist, hand, and finger) using X-ray images. A two-stage classification method based on MobileNet model is proposed. Enhanced X-ray image is fed into the first stage to detect bone type. Thereafter, the bone image is directed according to the result of the first stage to one of seven classifiers (one for each bone type) to detect the abnormality in the bone. MURA dataset is utilized as a performance dataset and average accuracy

Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

73.42% has been achieved after merging the two classification stages."

Ananthu et al. [25] employed in many medical imaging "applications for the diagnosis of diseases. However, one of the key issues is the limited availability of microscopic images for training the models. To overcome this difficulty, transfer learning techniques are put forward, researchers present a comparative analysis of different transfer learning models like MobileNet to detect acute lymphocytic leukemia (ALL) from blood smear cells. All models were trained on ALL-IDB2 dataset and achieved an accuracy of 97.88%," from MobileNet model.

Wang et al. [28] tried in their research "to improve and speed up bone age assessments by using different object detection methods to detect and segment bones anatomically important for the assessment and using these segmented bones to train deep learning models to predict bone age. A dataset consisting of 12811 X-ray hand images of persons ranging from infant age to 19 years of age was used. In the first research question, compared the performance of three state-of-the-art object detection models: Mask R-CNN, Yolo, and RetinaNet, was selected the best performing model, Yolo, to segment all the growth plates in the phalanges of the dataset, proceeded to train four different pre-trained models: Xception, InceptionV3, VGG19, and ResNet152, using both the segmented and unsegmented dataset and compared the performance, achieved good results using both the unsegmented and segmented dataset, although the performance was slightly better using the unsegmented dataset. The analysis suggests that we might be able to achieve a higher accuracy using the segmented dataset by adding the detection of growth plates from the carpal bones, epiphysis, and the diaphysis. The best performing model was Xception, which achieved a mean average error of 1.007 years using the unsegmented dataset and 1.193 years using the segmented dataset."

5.1 Summary of the previous Studies

In the current study we have 14 classes of bone abnormalities. Thus the current study is a multi-label classification.

Table 1 presents a summary of the most important previous studies that were discussed, where the reference, machine learning methodology, detailed information about the dataset used, the programming language used, if any, and the best accuracy obtained.

Table 1: Summarizes of the most important discussed previous studies

Study Reference	Methods used	Dataset Attributes	Best Accuracy
(El-Saadawy, Hadeer, et al., 2020) [24]	MobileNet	X-ray images	73.42
(Ananthu, K. S., et al., 2021) [25]	MobileNet	images	97.88
(Wang, Shudong, Dong, Liyuan, Wang, Xun and Wang, Xingguang., 2020) [28]	Xception, InceptionV3, VGG19, and ResNet152	CT images	83.71
(Russakovsky, O. et al., 2015) [13]	InceptionV3	Images	79.70
(Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen & H. Greenspan., 2015) [17]	MobileNet	Images	94.00
(Dimitris, K., Ergina, K., 2017) [29]	MobileNet	Biomedical images	75.83
(N. Dhungel, G. Carneiro and A. P. Bradley., 2017) [30]	Xception	Mammogram images	79.00
(Moran, Maira, et al., 2021) [32]	Xception	images	73.30
(Westerberg, Erik., 2020) [33]	CNN	X-ray images	97.40
(Ding, Y., et al. (2019) [35]	CNN	images	98.00
(Arun D. Kulkarni 2022) [36]	VoxCNN, RNN	MRI images	88.00
Esteva, A., Chou, K., Yeung, S. et al. (2021[19]	InceptionV3	Clinical images	72.10
(Cernazanu-Glavan, C., & Holban, S., 2013)[20]	Resnet50	X-ray images	80.25
(Sitaula, C., Hossain, M.B., 2021) [21]	VGG16, VGG19	X-ray images	79.58
(Esteva, A., Kuprel, B. et al. 2017) [22]	VGG16	Clinical images	72.10

Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(Gulshan V, Peng L, Coram M, et al., 2016) [11]	InceptionV3	images	98.50
(Moreno-García, Carlos, et al., 2020) [15]	VGG16, Resnet50, InceptionV3	X-ray images	82.70
(MORAN, Maira Beatriz Hernandez, et al., 2000)[23]	ResNet and Inception	images	92.30

6. METHODOLOGY

The methodology that we used in this study has the following steps: dataset collection, data prep-processing, feature engineering, data splitting, proposed deep learning model, model training, and testing as in Figure 1.

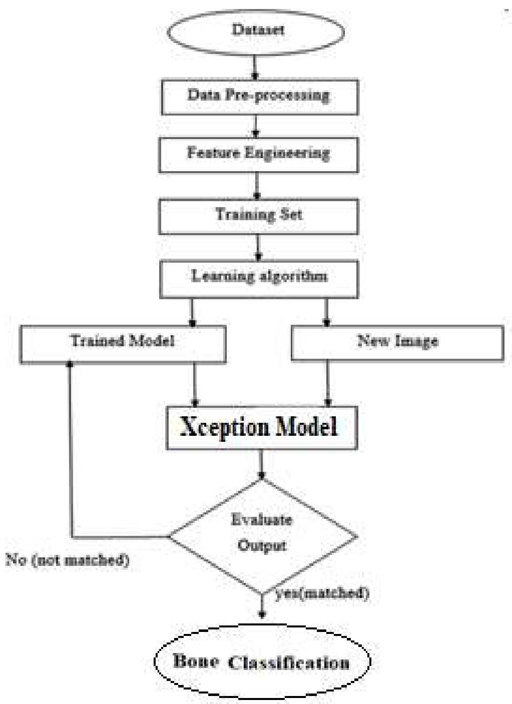

Figure 1: methodology flowchart

5.1 Data Collection

We have collected the Bone Abnormalities Images from Kaggle depository. The dataset is called Mura-v1.1. The Mura-v1.1 contains 42,000 X-rays images of abnormal bones. The data set has 14 classes: Elbow-Pos, Finger-Pos, Forearm-Pos, Hand-Pos, Humerus-Pos, Humerus-Neg, Shoulder-Pos, Wrist-Pos, Finger-Neg, Elbow-Neg, Forearm-Neg, Shoulder-Neg, Wrist-Neg, Hand-Neg.

5.2 Data Splitting

We have split the dataset into three datasets: training and testing datasets. The ratio of splitting is 80%, 20% respectively. Furthermore, the Training dataset was split into train and valid datasets with 60% for train and 20% for valid datasets. The new total number of X-Ray images is 42,000. Each class of the 14 classes has 3,000 images. Samples of the 14 classes are shown in Figure 2.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

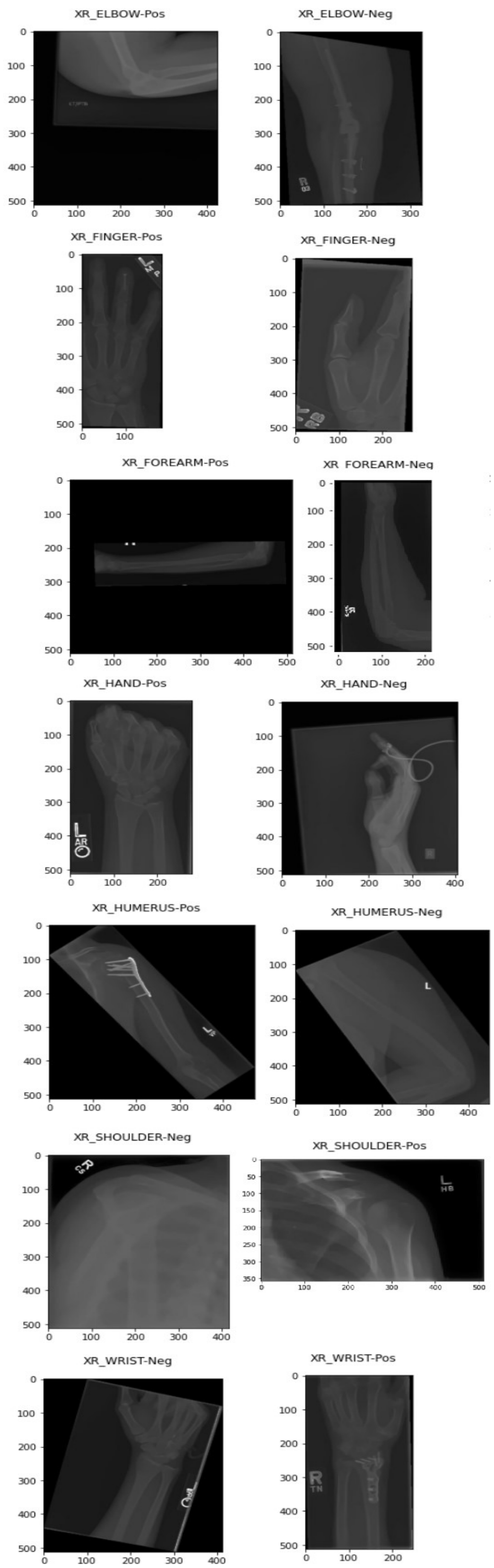

Figure 2: Samples of the Bone X-ray images dataset

5.3 Performance measures

We used the most common criterion for measuring the performance of the proposed Xception model:

- Precision is defined by True Positive divided by the summation of True Positive and False Positive as in equation 1.
- Recall is defined by True Positive divided by the summation of True Positive and False Negatives as in equation 2.
- F1-score is defined by 2 times Precision times Recall divide by the summation of Precision and Recall as in equation 3.

$$P = \frac{True\ positives}{True\ positives\ + False\ positives} \quad \text{... eq(1)}$$

$$R = \frac{True\ positives}{True\ positives\ + False\ negatives} \quad \text{... eq(2)}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \text{............ eq(3)}$$

5.4 Proposed model

In the current study we proposed to utilize Xception model for the classification of 14 classes of bone abnormalities of X-ray Images. The original architecture of the VGG16 before modification is shown in Figure 3. The original Xception model was used to classify 1000 classes of different objects. The original Xception model cannot be used directly to classify the 14 classes of bone abnormalities. Therefore, we need to modify it by replacing the top layer (classifier) with our own classifier. The modified Xception model is represented in Figure 4.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

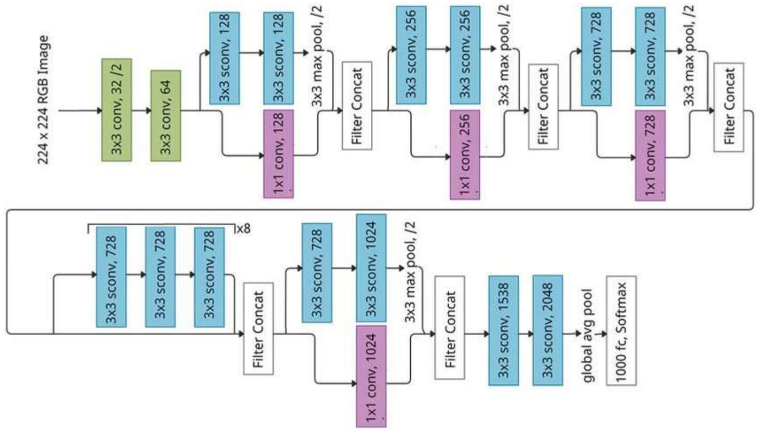

Figure 3: Original Xception model

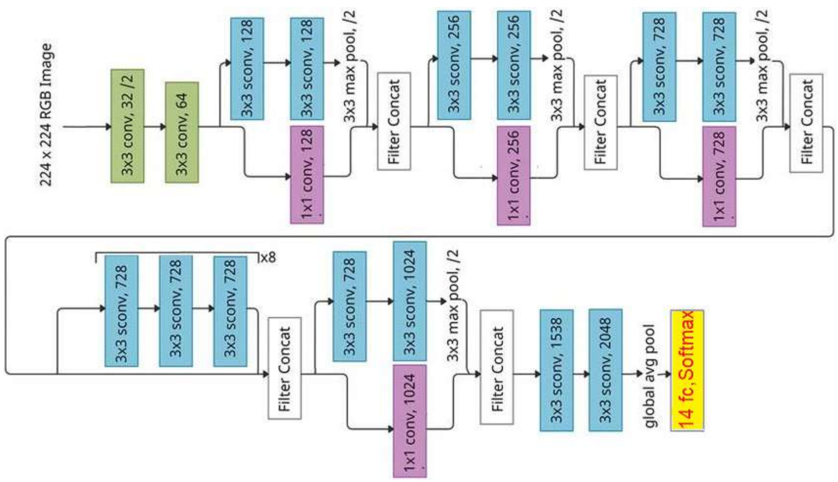

Figure 4: Modified Xception model

3.5 Model Training and Validating and Testing

The proposed Xception model was trained using the train dataset and validated using the valid dataset. The training was done through 200 epochs with learning rate (0.0001), batch size (256) and Softmax function and Adam as Optimizer. Furthermore, to overcome the training problems that can occur during the training, augmentation technique was utilized. Figure 5 and Figure 6 shows the loss and accuracy of the training and validation of the proposed Xception model. After finishing the training of the proposed Xception model, we tested it using the testing dataset.

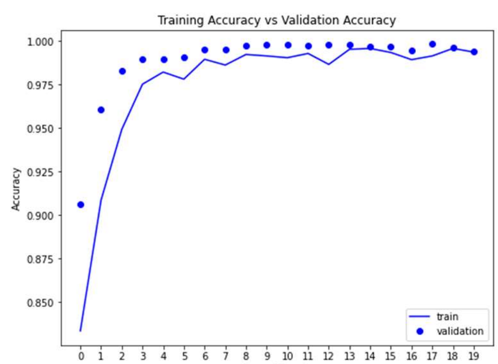

Figure 5: Training Accuracy vs. Validation Accuracy

Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

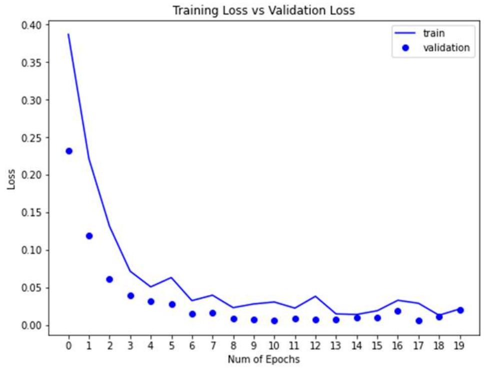

Figure 6: Training loss vs. Validation loss

7. RESULTS AND DISCUSSION

The proposed Xception model attained Training Accuracy (99.82%), Validating Accuracy (99.78%) and Testing Accuracy (85.13%). In terms of loss, in the customized model the training Loss is (0.056), Validating Loss is (0.0063), Testing Loss is (0.081). In terms of the time required for training and testing, the proposed Xception model required 5100 seconds for training and 15.7 seconds for testing.

Table 2 shows the precision, Recall, F1-Score of each class in the dataset in terms of the 14 classes that the proposed Xception model used for the classification of bone abnormalities: Elbow-Pos, Finger-Pos, Forearm-Pos, Hand-Pos, Humerus-Pos, Humerus-Neg, Shoulder-Pos, Wrist-Pos, Finger-Neg, Elbow-Neg, Forearm-Neg, Shoulder-Neg, Wrist-Neg, Hand-Neg. The proposed Xception model attained an average Precision (85.20%), Recall (85.13%) and F1-Score (85.07%). Furthermore, the ROC Curve measure for each class in the dataset reached 99%.

Table 2: Xception Precision, Recall, and F1-Score of each class in the dataset

Class	Precision	Recall	F1-score	Number of images used
XR_ELBOW-Pos	0.8639	0.8624	0.8631	574
XR_ELBOW-Neg	0.8730	0.8624	0.8677	574
XR_FINGER-Pos	0.8319	0.8711	0.8511	574
XR_FINGER-Neg	0.8407	0.7997	0.8196	574
XR_FOREARM-Pos	0.9592	0.9826	0.9707	574
XR_FOREARM-Neg	0.7981	0.8606	0.8282	574
XR_HAND-Pos	0.8504	0.7822	0.8149	574
XR_HAND-Neg	0.9793	0.9878	0.9835	574
XR_HUMERUS-Pos	0.7562	0.6376	0.6919	574
XR_HUMERUS-Neg	0.6927	0.7920	0.7390	572
XR_SHOULDER-Pos	0.9828	0.9930	0.9879	574
XR_SHOULDER-Neg	0.7413	0.7787	0.7596	574
XR_WRIST-Pos	0.7901	0.7474	0.7681	574
XR_WRIST-Neg	0.9684	0.9599	0.9641	574
Accuracy			0.8513	8034
macro avg	0.8520	0.8512	0.8507	8034
weighted avg	0.8520	0.8513	0.8507	8034

The current result cannot be compared with previous studies for two reasons:

- All previous studies use 2 classes (binary classification), they test whether there is a bone abnormality or not. In the current study, we have 14 classes of bone abnormalities. We classify the image as (Elbow, Finger, Forearm, Hand, Humerus, Shoulder, and Wrist) and detect whether that part has bone abnormalities or not.
- The dataset somewhat is larger than they used in the previous studies.

8. CONCLUSION

Bones are both strong and flexible tissue made up of collagen and calcium phosphate. They mainly contribute to the movement of the human body and also serve as a protective shield for the body's soft organs such as the heart, lungs, and brain. Without them, the human body would not be built to function normally. But sometimes, due to various accidents, a person is exposed to some diseases such as infection or injury, which lead to defects in normal growth and shape of bone structure.

The main aim of the study is to propose a deep learning model for the classification of the 14 classes of bone abnormalities. Xception model was proposed to do the job. We modified the Xception model to suite the 14 classes that we have of bone abnormalities.

The Dataset was collected from Kaggle and boosted using data augmentation. We split the dataset into two datasets: training and testing. We trained, validated, and tested the modified Xception model.

The proposed Xception mode attained Precision (85.96%), Recall (85.82%) and F1-Score (85.77%).

9. FUTURE WORK

We are planning to continue to fine tune our proposed model to get better accurcey, F1-score, Recall, Precision, time performance. In this way we contribute in helping specialist in diagnosing upper bone abnormalites quickly and efficiently.

Author Contribution Statement

The authors confirm contribution to the paper as follows: study conception and design: AB, SA; data collection: AB, MA, SA; analysis and interpretation of results: AB. Author; draft manuscript preparation: AB, SA, MA. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

[1] Abunasser, B. S., et al., "Breast Cancer Detection and Classification using Deep Learning Xception Algorithm" International Journal of Advanced Computer Science and Applications(IJACSA), vol. 13, no. 7, pp. 223-228, 2022. http://dx.doi.org/10.14569/IJACSA.2022.0130729.

[2] Obaid, T., et al. "Factors Contributing to an Effective E- Government Adoption in Palestine" Lecture Notes on Data Engineering and Communications Technologies, 127, pp. 663–676, 2022.

[3] Abunasser, B. S., et al., "Prediction of Instructor Performance using Machine and Deep Learning Techniques" International Journal of Advanced Computer Science and Applications(IJACSA), vol. 13, no. 7, pp. 78-83, 2022. http://dx.doi.org/10.14569/IJACSA.2022.0130711.

[4] Saleh, A., et al. "Brain tumor classification using deep learning" Proceedings - 2020 International Conference on Assistive and Rehabilitation Technologies, iCareTech 2020, 2020, pp. 131–136, 9328072.

[5] Arqawi, S., et al. "Integration of the dimensions of computerized health information systems and their role in improving administrative performance in Al-Shifa medical complex" Journal of Theoretical and Applied Information Technology, vol. 98, no. 6, pp. 1087–1119, 2020.

[6] Elzamly, A., et al. "Assessment risks for managing software planning processes in information technology systems" International Journal of Advanced Science and Technology, vol. 28, no. 1, pp. 327–338, 2019.

[7] Albatish, I.M., et al. Modeling and controlling smart traffic light system using a rule based system. Proceedings - 2019 International Conference on Promising Electronic Technologies, ICPET 2019, pp. 55–60, 2019, 8925318

[8] Elzamly, A., et al. "A new conceptual framework modelling for cloud computing risk management in banking organizations" International Journal of Grid and Distributed Computing, vol. 9, no. 9, pp. 137–154, 2016.

[9] Mady, S.A., et al. "Lean manufacturing dimensions and its relationship in promoting the improvement of production processes in industrial companies" International Journal on Emerging Technologies, vol. 11, no. 3, pp. 881–896, 2020,

[10] Abu Ghosh, M.M., et al. "Secure mobile cloud computing for sensitive data: Teacher services for Palestinian higher education institutions" International Journal of Grid and Distributed Computing, vol. 9, no. 2, pp. 17–22, 2016.

[11] Gulshan, V., & Peng, L. et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA, vol. 316, no. 22, pp. 2402-2420, 2016.

[12] Buhisi, N. I., et al. "Dynamic programming as a tool of decision supporting" Journal of Applied Sciences Research, vol. 5, no. 6, pp. 671-676, 2009.
[13] Russakovsky, O., & Deng, J. et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.
[14] Naser, S. S. A. "Developing an intelligent tutoring system for students learning to program in C++" Information Technology Journal, vol. 7, no. 7, pp. 1051-1060, 2008.
[15] Moreno-García, Carlos, et al. Assessing the clinicians pathway to embed artificial intelligence for assisted diagnostics of fracture detection. CEUR Workshop Proceedings, 2020.
[16] Naser, S. S. A. "Developing visualization tool for teaching AI searching algorithms" Information Technology Journal, vol. 7, no. 2, pp. 350-355, 2008.
[17] Bar, Y., & Diamant, I. et al. Chest pathology detection using deep learning with non-medical training. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp. 294–297, 2015.
[18] Fawzy A., et al. "Mechanical Reconfigurable Microstrip Antenna" International Journal of Microwave and Optical Technology, vol. 11, no. 3, pp.153-160, 2016.
[19] Naser, S. S. A. "Intelligent tutoring system for teaching database to sophomore students in Gaza and its effect on their performance" Information Technology Journal, vol. 5, no. 5, pp. 916-922, 2006.
[20] CERNAZANU-GLAVAN, C., & HOLBAN, S. Segmentation of Bone Structure in X-ray Images using Convolutional Neural Network. Advances in Electrical and Computer Engineering, vol. 13, no. 1, pp. 87–94, 2013.
[21] Sitaula, C., & Hossain, M. B. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. Applied Intelligence, vol. 51, no. 5, pp. 2850-2863, 2021.
[22] Esteva, A., Kuprel, B. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature, vol. 542, no. 7639, pp. 115–118, 2017.
[23] Moran, M. B. H., & Faria, M. et al. On using convolutional neural networks to classify periodontal bone destruction in periapical radiographs. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2036–2039, 2020.
[24] El-Saadawy, H., & Tantawi, M. et al. A Two-Stage Method for Bone X-Rays Abnormality Detection Using MobileNet Network. Advances in Intell,gent Systems and Computing, pp. 372–380, 2020.
[25] Ananthu, K. S., & Krishna Prasad, P. et al. Acute Lymphoblastic Leukemia Detection Using Transfer Learning Techniques. Intelligent Sustainable Systems, pp. 679–692, 2021.
[26] Naser, S. S. A. "JEE-Tutor: An intelligent tutoring system for java expressions evaluation" Information Technology Journal, vol. 7, no. 3, pp. 528-532, 2008.
[27] Jawad Yousef Ibrahim Alzamily, Syaiba Balqish Ariffin, Samy S. Abu-Naser. "Classification of Encrypted Images Using Deep Learning –ResNet50." Journal of Theoretical and Applied Information Technology, 100, no. 20, 2022.
[28] Wang, S., & Dong, L. Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy. Open Medicine, vol. 15, no. 1, pp. 190–197, 2020.
[29] Dimitris, K., Ergina, K. Concept detection on medical images using Deep Residual Learning Network. In working notes of conference and labs of the evaluation forum. Springer, 2017.
[30] Dhungel, N., Carneiro, G., & Bradley, A. P. Fully automated classification of mammograms using deep residual neural networks. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 310–314, 2017.
[31] Abu-Naser, S.S., et al. "An expert system for endocrine diagnosis and treatments using JESS" Journal of Artificial Intelligence, vol. 3, no .4, pp. 239-251, 2010.
[32] Moran, M., & Faria, M. et al. Classification of Approximal Caries in Bitewing Radiographs Using Convolutional Neural Networks. Sensors, vol. 21, no. 15, pp. 5192-5199, 2021.
[33] Westerberg, E. AI-based Age Estimation using X-ray Hand Images . A comparison of Object Detection and Deep Learning models, 2020.
[34] Abu Naser, S.S. "Evaluating the effectiveness of the CPP-Tutor, an intelligent tutoring system for students learning to program in C++" Journal of Applied Sciences Research, vol. 5, no. 1, pp. 109-114, 2009.
[35] Ding, Y., & Sohn, J. H. et al. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain.

Radiology, vol. 290, no. 2, pp. 456–464., 2019.
[36] Arun D. Kulkarni, "Deep Convolution Neural Networks for Image Classification" International Journal of Advanced Computer Science and Applications, vol. 13, no. 6, 2022.
[37] Abdelbaset R. Almasri , Nor Adnan Yahaya, and Samy S. Abu-Naser. "Instructor Performance Modeling For Predicting Student Satisfaction Using Machine Learning-Preliminary Results." Journal of Theoretical and Applied Information Technology 100, no. 19, pp. 5481-5496, 2022.
[38] Alaa M. A. Barhoom, Mohammed Rasheed J. Al-Hiealy, Samy S. Abu-Naser. "Bone Abnormalities Detection and Classification Using Deep Learning-Vgg16 Algorithm." Journal of Theoretical and Applied Information Technology, 100, no. 20, 2022.
[39] Samer M. Arqawi, Mohammed A. Abu Rumman, Eman Akef Zitawi, Anees Husni Rabaya, Ahmad Saleh Sadaqa, Basem S. Abunasser, Samy S. Abu-Naser. "Predicting Employee Attrition and Performance Using Deep Learning." Journal of Theoretical and Applied Information Technology, 100, no. 20, 2022.