IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⅔ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

BONE AGE ASSESSMENT USING MACHINE LEARNING AND IMAGE PROCESSING

Tushar Jain[1], Aftab Khan[2]

[1,2]Department of Computer Engineering, M.H. Saboo Siddik College of Engineering, Mumbai, India

Abstract: Bone age assessment and its comparison with the chronological age is a crucial task to determine the disorders and their effects on the bone when there are fewer documents. It is a time-consuming activity that is performed by the doctors by the method known as ossification. It can be automated with machine learning techniques. In the proposed system, the images of hand radiographs are preprocessed using data augmentation and the feature extraction is done using pre-trained Mobilenet and Xception models. The obtained results have shown that the Xception model gives the best MAE as compared with Mobilenet.

Keywords: Bone Age Assessment, X-ray images, Xception, Mobilenet, Transfer Learning, Deep Learning.

1. INTRODUCTION

Bone age assessment is the process in which the age of bone is given with the help of various techniques available to doctors. It is mostly done to check the maturity of the bone and also, if there is any problem in the growth then it can be resolved. These are performed on children as they have a growth rate more than others but nowadays old people are also required to assess their bones as they have many problems related to it. Bone age is related to skeletal maturity and also estimation of age is based on an individual's skeletal maturity whereas the person's age is calculated by considering the date of birth. The significance of surveying a person's skeletal development or bone age and its correlation with their ordered age emerges from two primary points: First, from a clinical point, the appraisal of bone age is helpful for the conclusion and treatment of pediatric endocrinology, orthodontics, and pediatric muscular issues, as well as additionally being considered in assessments of a singular's last tallness. Second, from a lawful angle, bone age evaluation is significant to distinguish whether an individual is what is going on where confirmed archives are deficient [1].
Radiologists use a process known as ossification to detect bone age. Hardening in bone redesigning is the most common way of setting down new bone material by cells called osteoblasts. It is inseparable from bone tissue arrangement.

Two cycles are bringing about the arrangement of ordinary, solid bone tissue:
Intramembranous hardening is the immediate setting down of bone into the crude connective tissue, endochondral solidification includes ligament as an antecedent. Heterotopic hardening is an interaction bringing about the arrangement of bone tissue that is regularly abnormal, in an extraskeletal area. Calcification is frequently mistaken for solidification. Calcification is inseparable from the development of calcium-based salts and gems inside cells and tissue. A cycle happens during solidification, however not the other way around [2]. Current methods are based on various techniques of machine learning and image processing where they use hand radiographs to predict maturity. The meaning of ML is wide. ML is the investigation of apparatuses and strategies for distinguishing designs in information. These examples can then be utilized to either build how we might interpret the current world or make forecasts about the future.ML draws on ideas from many fields including software engineering, measurements, and enhancement. At their center, practically everything ML issues can be formed as an improvement issue concerning a dataset. In such settings, the objective is to observe a model that best makes sense of the information. While there are various sorts of ML, most applications can be categorized as 1 of 3 classes: supervised, unsupervised, or reinforcement learning [3].
Here supervised learning is used as the radiographs of patients are taken into consideration and age is predicted from that. Deep learning and transfer learning is used for effectively carrying out feature extraction without segmentation so it is very useful in our case. The proposed system consists xception with a sequential model to train the model which gave us very good results.
Section 2 deals with the Literature survey. Section 3 deals with the Proposed System. Section 4 deals with the dataset. Section 5 contains the Implementation and Experimental Results. Section 6 is about Conclusion and Future work.

2. LITERATURE SURVEY

Shaower li have designed the system to assess bone maturity they have made the system using various machine learning and deep learning algorithms which have helped in finding the maturity of the bone efficiently [8].

DOI: 10.17148/IJARCCE.2022.11450

The advantages of the system are that the accuracy of the system is good. The system has good prediction performance. The precision of the proposed system is good.The proposed system has a fully automated approach.
The disadvantages of the system are that there is a lack of medical experience in the system. The system does not consider gender information. It works on small datasets. Size normalization is not done.
Almost all the research papers we analyzed were working on the RSNA dataset. Table 1 illustrates a comparative representation of their works:

Table 1: Comparison between related works

Algorithms/ Methods Used	Method of Evaluation	Remarks	Future Scope	Gaps
TW3, Convex-Hull method, RANSAC	Finding 13 ROI's, Finding Big ROI's	no actual evaluation, just methods for finding Big ROI's were discussed	Automation, need to evaluate actual assessment	need to find all the 13 ROI's, no actual evaluation or assessment[1]
CNN, Image Compression, SE-ResNet, Improved loss method, Keras for deep learning	Lossless image compression to reduce time, SE-ResNet to extract features, improved loss function uses regression to evaluate bone age	Outperforms others in terms of time, error and huge variability in input images	Compression and ResNet modules can be used for recognition in various medical problems	More time is required for larger images[4]
U-Net, Active Learning(AL), 6 CNN modules, Oracle for annotation, Ensemble Regression Algorithms	U-Net module is used to produce hand masks, AL is implemented to handle annotation, CNN modules are used to extract features, Ensemble Algorithms are used for BAA	Outperformed others in MAE	The proposed framework should be validated on other medical decision problems, the effectiveness of the AL framework needs to be proved	Implementation is long and requires specific framework modules[5]
Unsupervised Learning, Image processing pipeline, MLP, CNN	unsupervised learning is used to locate informative region to reduce annotation, MLP was reduced for prediction, gender information was also used for prediction	Outperformed others in terms of accuracy and time	A deep learning framework can be proposed to extract more fine-grained features and automation can also be achieved	Implementation is time consuming and tough[6]
CNN, Deep Learning(Keras), DeepLab V3 plus, MobileNet V1, Xception Regressor	Image registration is used for segmenting hand and locating key points, then Xception regressor network is used for BAA	Outperforms its deep learning competitors in MAE, MSE. It considers 20,000,000 parameters for evaluation	Addition of convolution filters can improve the work	The work does not take gender into consideration, which can affect the accuracy[7]
DT, KNN, LD, SVM	Accuracy, Specificity & precision performance metrics	LD classifier failed with ResNet based features due to limited computing resource	To evaluate the proposed system using larger dataset and classification models	Less classification model used with small dataset[8]

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022
DOI: 10.17148/IJARCCE.2022.11450

CNN, DNN, DBN, RNN	Efficient Performance(CNN)	Need for modification in architecture	Structured and text free reports can be used for training of the model	Slicing of input image is done in fixed kernel size but variable size can be taken[9]
CNN, Regressor Network, AxNet	Improved Prediction Performance	Gender Information is Omitted	Size normalization can be done, Multi-Scale capability	Gender is not taken into consideration[10]
CNN, ANN, SVM, Bayesian Networks, Decision Trees, KNN	Accuracy and Fully automatic	Left-Hand radiographs are manually analyzed using GP or TW method as reference data	Comparative study between radiography and MRI based age estimation solution is needed	Lack of a method to confirm that determined age represents true bone age[11]
CNN, ANN, Regression methods, SVM, Bayesian Networks, Decision Trees, KNN	Fully Automated Approach	Lack of medical experience	Investigation of other ROI for BAA	Information from medical experts[12]

The system reviewed has the following limitations.
The system is not able to assess the bone maturity for adults. The system is assessing the maturity of the bone only for children below 17 years and also in that it is time-consuming. Need to find all the 13 ROI's, no actual evaluation or assessment. More time is required for larger images. Implementation is long and requires specific framework modules. Implementation is time-consuming and tough. The work does not take gender into consideration, which can affect the accuracy. A Smaller classification model was used with a small dataset. Slicing of the input image is done in fixed kernel size but variable size can be taken. Gender is not taken into consideration. Method to confirm that determining age represents true bone age. Information from medical experts.
This work recommends a system that will implement bone age assessment using Transfer Learning, Deep Learning, and Regression methods that will help in assessing the maturity of the bone.
This study covers the survey of different bone maturity assessment methods using machine learning and deep learning. This study analyzes to get information from medical experts. This study analyzes gender into consideration while training.

3. PROPOSED SYSTEM

In this section, bone maturity assessment is proposed based on hand X-ray images. It aims to determine the age of the person using the hand radiographs. Figure 1 shows the steps of the system. These are the following steps: first the input image is taken from the user and then the most important step is image preprocessing in which the image is resized and the image is taken as RGB. The next step is to train the model which is done by using CNNs Xception at last the user gets the result which is the age. In the system, we have calculated the mean bone age and the standard bone age from the dataset, and from that, we have calculated the z score. A z-score estimates the distance between an informative item and the mean utilizing standard deviations. Z-scores can be positive or negative. The sign lets you know whether the perception is above or underneath the mean. For instance, a z-score of +2 demonstrates that the information point falls two standard deviations over the mean, while a - 2 connotes it is two standard deviations beneath the mean. A z-score of zero corresponds to the mean. Analysts likewise allude to z-scores as standard scores. It is calculated to assess how an individual's bone age compares with the mean population bone age. It helped in the learning process very nicely.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

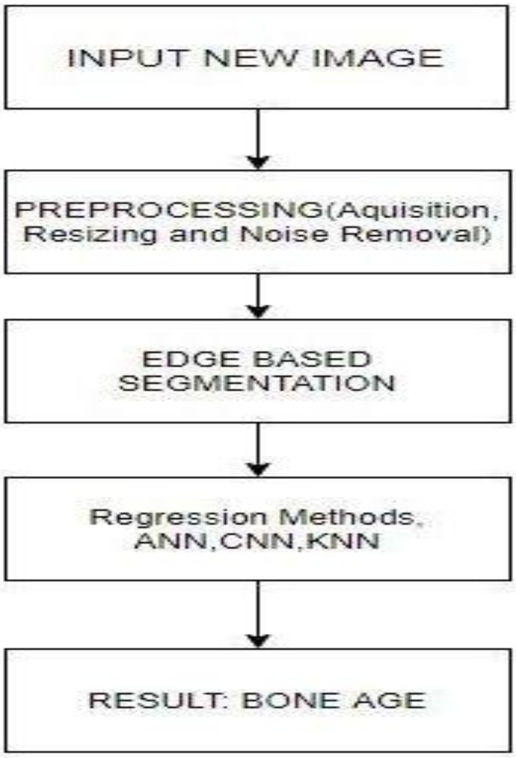

Figure1: Block diagram of the proposed bone age assessment system

For example, the distribution of bone age for men is normally distributed with a mean of about 80 and a standard deviation of 20. If a certain man has a bone age of 100, we would calculate their z-score to be:

$z = (x - \mu) / \sigma$
$z = (100 - 80) / 20$
$z = 1$

The second thing is that we also have taken custom metrics to judge the performance of the model. A metric function is used to have a proper judgment of the model.
Metric functions are the same as loss functions, but the difference is that the results from assessing a metric are not utilized when training the model. Note that you can utilize any loss function as a metric. Much like loss functions, any callable with the signature metric that returns an array of losses can be passed to the compile method as a metric.

Image Preprocessing
In this operation, the hand X-ray images are ready for the next step, which is converting the image to RGB image and then resizing them to 256*256, and then scaling them 1/255.0 all of this is done with Keras preprocess input function. For this data augmentation is performed. Data Augmentation is a bunch of methods to misleadingly expand how much information is by creating new items from existing information. This incorporates rolling out little improvements to information or utilizing profound learning models to produce new items. In this process making simple changes in the images is done and it is very popular. Many operations are performed in this such as padding, random-rotating, re-scaling, vertical and horizontal flipping, translation, cropping, zooming, color modification, and grayscaling. Advantages of data augmentation are further developing model forecast precision, adding additional preparation information into the models, forestalling information shortage for better models, lessening information overfitting (for example a mistake in measurements, implies a capacity related to near a restricted arrangement of informative items) and making fluctuation in information, expanding speculation capacity of the models, assisting resolve with classing unevenness issues in characterization, Diminishing expenses of gathering and marking information, Empowers interesting occasion forecast and Forestalls information security issues. The data augmentation step is performed to build a large dataset that is suitable for deep neural networks and the methods used here are shuffling and flipping it vertically.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

Feature Extraction

Convolutional Neural Network is a network that is very useful and has achieved many breakthroughs in the field of deep learning, machine learning, and computer vision. CNN can complete the task without being affected by titling, translation, and scaling; they have 3 layers: the convolutional layer, pooling layer, and fully connected layer.

The convolutional layer's job is to compute the weighted sum, add the bias value to it, and apply an activation function called the rectifier linear unit (ReLu) to the addition result, which is described using Eq. 1. Pooling layers, on the other hand, are used to control over-fitting by reducing the number of features obtained from the convolutional layer. Finally, the completely linked layers strive to collect all of the descriptor features that will be categorized using the final layer.

$$Relu(x) = max(0, x) \qquad (1)$$

In the proposed system two of the pre-trained convolutional neural networks are Mobilenet and Xception. Using the notion of transfer learning, the feature extraction stage is performed. Transfer learning is commonly employed for a variety of purposes. Due to the lack of big datasets, training a CNN from scratch with random beginning values is challenging. As a result, utilizing the weights of a pre-trained net as beginning values can help solve a lot of issues. Second, training a highly deep network from the start is a time-consuming operation that necessitates the use of expensive GPUs and complex processors. Finally, there is no clear theoretical guidance on how to choose the best topology, training technique, parameter values, and so forth [12].

MobileNet is distinguished by its use of depth-wise separable convolutions, which may be thought of as conventional convolutions divided into depthwise and 1 by 1 pointwise convolution. Each input channel is filtered independently in depthwise convolution. The depthwise convolution outputs are then linearly combined using 1 by 1 pointwise convolution. This factorization has resulted in a significant decrease in model size and calculation cost. When compared to using ordinary convolutions, the calculation cost is lowered by around 8 to 9 times. The network has two hyperparameters: 1. width multiplier and 2. resolution multiplier. The width multiplier's purpose is to evenly thin the network at each layer by lowering the number of filters, whereas the resolution multiplier may be used to reduce picture resolution. We can get some smaller models by changing the width multiplier and resolution multiplier numbers, but this invariably results in a drop in overall accuracy. After each convolution layer in the model, batch normalization and the non–linear activation function ReLU were utilized. Downsampling is accomplished by employing stridden convolutions in both the first and depthwise convolution layers. After that, an Average Pooling layer, a Fully Connected layer, and a Softmax classifier are added. If the Depthwise convolution and Pointwise convolution layers are analyzed independently, the baseline MobileNet has 28 layers. The MobileNet v1 basic architecture depicted in Figure 2 is a schematic diagram with all of the layers [13].

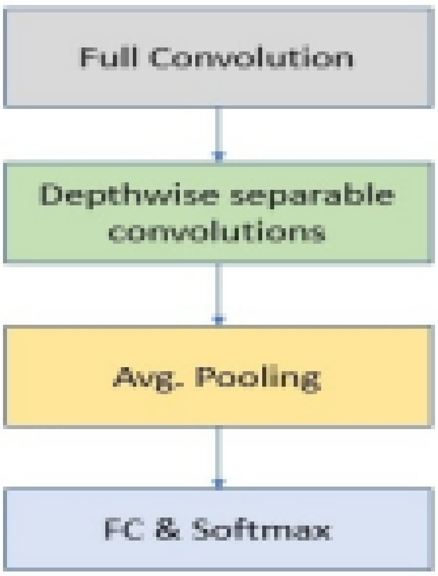

Figure 2: Mobilenet v1 baseline model [13].

In our system, the mobilenet layers are used with weights pre-trained on imagenet and it is trained with an x-ray dataset.Then we have taken global average pooling along with three dense layers with 1024, 1024, and 512 units along

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁙ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

with the activation function of relu. The last layer is the output layer which is dense with 1 unit as it contains our predicted value with activation function as linear activation function. We got the following graph as shown in Figure 3 which is produced after training and testing.

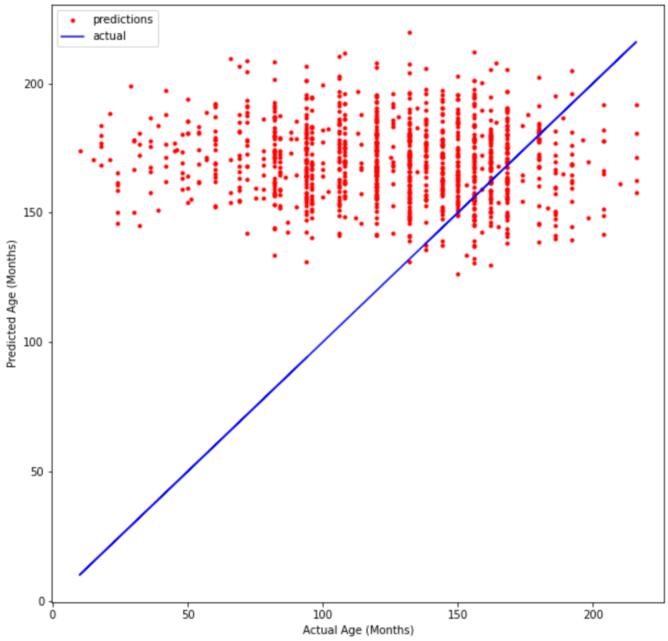

Figure 3: Graph of Mobilenet model on the RSNA bone age dataset.

The graph compares actual age to expected age, which is expressed in months.Figure 4 gives a detailed summary of the network's specs. The feature extraction basis of the Xception architecture is made up of 36 convolutional layers. We will just look at picture classification in our experimental assessment, therefore our convolutional base will be followed by a logistic regression layer. Optionally, fully-connected layers can be inserted before the logistic regression layer, as shown in the experimental assessment section. Except for the first and last modules, the 36 convolutional layers are divided into 14 modules, all of which contain linear residual connections surrounding them.

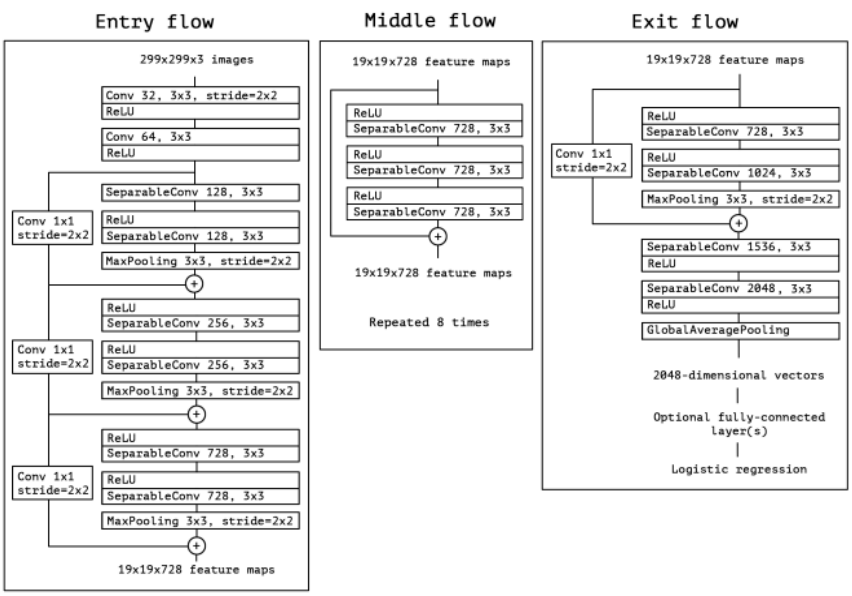

Figure 4: Xception architecture [14].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

In a nutshell, the Xception design is a depthwise separable convolution layer stack with residual connections. This makes the architecture incredibly simple to construct and change; using a high-level framework like Keras or TensorFlow-Slim, it only requires 30 to 40 lines of code, similar to an architecture like VGG-16, but unlike architectures like Inception V2 or V3, which are significantly more complicated to define. The Keras Applications module includes an open-source version of Xception that uses Keras and TensorFlow and is licensed under the MIT license [14].
Data augmentation is used to solve the problem of overfitting. The Xception model is not sequential hence at first a sequential layer is added with the Xception layer and the weights are taken from imagenet. Then we have not used the top layers and the weights are trained with our dataset. Now we added a global max-pooling layer, a flatten layer and then two dense layers with activation function relu and linear after training we have tested with a testing dataset and we got the following result shown in the graph in Figure 5.

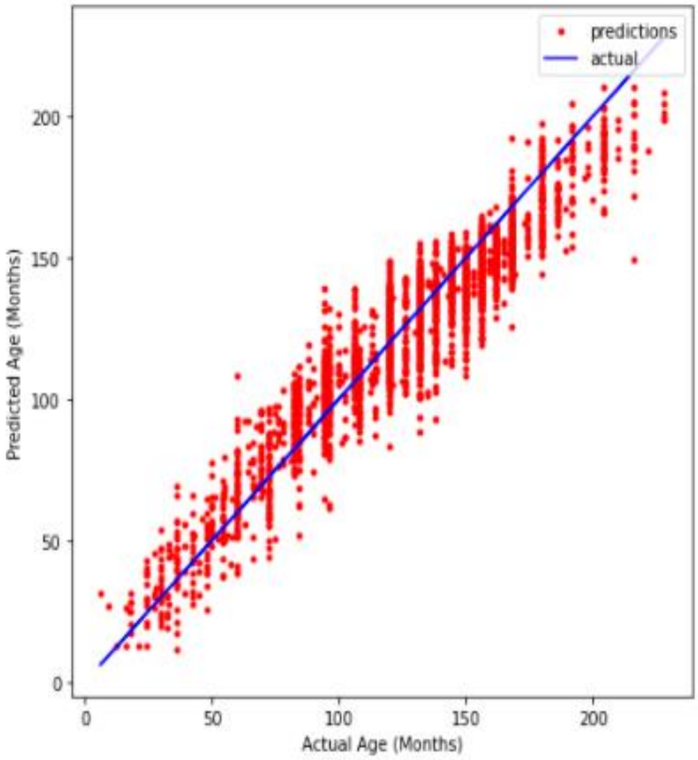

Figure 5: Graph of Xception model on the RSNA bone age dataset.

The graph shows actual age versus predicted age which is in months.

4. DATASET

The RSNA Bone Age dataset was utilized in this study. A total of 12000 hand X-ray pictures from people of all genders and ages are used to assess the proposed technology.

5. IMPLEMENTATION AND EXPERIMENTAL RESULTS

The proposed system of bone age assessment is done on google colab and Kaggle notebook also with the help of NumPy, pandas, matplotlib, TensorFlow, and Keras. The result is checked with mean absolute error(MAE) and mean squared error(MSE). As we have used both mobile net and xception. The difference between them is as follows: in the mobile net model we have not used the custom metrics to check the performance of the model instead we have used MAE and the loss function as MSE so the best value we got in MSE is 0.502 which is pretty good value. The graph is shown in Figure 4.

In the Xception model, we have taken custom metrics which is mean absolute error which is calculated in months to verify the difference in predicted age months with the real age in months and loss we have taken mean absolute error we got the loss in MAE which 0.09 which is a nice value and the difference in the months is only 5.12 months which is good in prediction. The graph of the result is shown in Figure 6. When both the systems were compared we got very good

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

results with the xception model and then the mobile net model. The following system can function and don't get affected by the outliers as data augmentation is done and transfer learning is used. Also, we have taken proper information from the doctors on the subject of bone age assessment. Following are the results obtained shown in Figures 6,7, 8, and 9.

Figure 6: The performance of the system using Xception

Figure 7: The performance of the system using Xception

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

Figure 8: The performance of the system using Xception

Figure 9: The performance of the system using Xception

6. CONCLUSION AND FUTURE WORK

In our current world, estimating human age has become a significant concern. Manual age estimate has several flaws. As a result, an efficient and effective automated human age estimate system is required. Transfer learning, on the other hand, is beneficial in a variety of machine learning and object identification problems. Using hand X-ray images, a transfer learning-based bone age assessment is proposed in this paper. The proposed system can predict bone age using hand radiographs. In the system the preprocess of the X-ray image is done and after that data augmentation is performed to overcome the problem of overfitting then feature extraction is done by mobilenet and xception Finally, prediction of bone age is done on the testing x-ray images. The xception based system gave very good results as compared to the mobilenet based system as in the mobilenet we have no used custom metrics whereas, in the xception it is used also the loss was

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11450

very less in the xception based system. The custom metrics performance was also good. As a result, promising performance in the area of autonomous human age estimate has been attained. We want to test the suggested system with a bigger dataset and various transfer learning models in the future.

REFERENCES

[1] Ana Luiza Dallora, P. A. (2019). Bone age assessment with various machine learning techniques: A systematic literature review and meta-analysis. Plosone, 22.
[2] Wikipedia. (2022, March 05). Ossification. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Ossification
[3] Shenoy, J. W. (2017). Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. HEALTHCARE EPIDEMIOLOGY, 5.
[4] Byoung-Dai Lee, M. S. (2020). Automated Bone Age Assessment Using Artificial Intelligence: The Future of Bone Age Assessment. Korean Journal of Radiology, 9.
[5] Cristina STOLOJESCU-CRIŞAN, Ş. H. (2021). A Comparison of X-Ray Image Segmentation Techniques. AECE, 8.
[6] JIN HE, D. J. (2020). Fully Automatic Model Based on SE-ResNet for Bone Age Assessment . IEEE Access, 7.
[7] Kelmara Arruda Pinho, P. F. (2021). Morphometric bone assessment of men and women in different age ranges by cone beam computerized tomography. Research, Society and Development, 11.
[8] Lee, C. L.-D. (2019). Deep-Learning based Finger Big ROIs Extraction for Bone Age Assessment in Smart Intelligence Systems. Journal of Engineering and Applied Sciences, 7.
[9] Mohd Asyraf Zulkifley, S. R. (2020). Automated Bone Age Assessment with Image Registration Using Hand X-ray Images. Applied Sciences, 16.
[10] Muhammad Waqas Nadeem, H. G. (2020). Bone Age Assessment Empowered with Deep Learning: A Survey, Open Research Challenges and Future Directions. Diagnostics, 22.
[11] NOOR MUALLA, E. H. (2020). AUTOMATIC BONE AGE ASSESSMENT USING HAND X-RAY IMAGES. Journal of Theoretical and Applied Information Technology, 12.
[12] Xiaoying Pan, Y. Z. (2020). Fully Automated Bone Age Assessment on Large-Scale Hand X-Ray Dataset. Hindawi, 12.
[13] El-Sharkawy, D. S. (2019). Thin MobileNet: An Enhanced MobileNet Architecture. IEEE, 6.
[14] Chollet, F. (n.d.). Xception: Deep Learning with Depthwise Separable Convolutions. CVPR, 8.