



The Open Biomedical Engineering Journal

Content list available at: <https://openbiomedicalengineeringjournal.com>



RESEARCH ARTICLE

Identifying Skeletal Maturity from X-rays using Deep Neural Networks

Suprava Patnaik^{1,*}, Sourodip Ghosh¹, Richik Ghosh¹ and Shreya Sahay¹

¹Department of Electronics Engineering, KIIT University, Bhubaneswar, India

Abstract: Skeletal maturity estimation is routinely evaluated by pediatrics and radiologists to assess growth and hormonal disorders. Methods integrated with regression techniques are incompatible with low-resolution digital samples and generate bias, when the evaluation protocols are implemented for feature assessment on coarse X-Ray hand images. This paper proposes a comparative analysis between two deep neural network architectures, with the base models such as Inception-ResNet-V2 and Xception-pre-trained networks. Based on 12,611 hand X-Ray images of RSNA Bone Age database, Inception-ResNet-V2 and Xception models have achieved R-Squared value of 0.935 and 0.942 respectively. Further, in the same order, the MAE accomplished by the two models are 12.583 and 13.299 respectively, when subjected to very few training instances with negligible chances of overfitting.

Keywords: Bone age identification, RSNA Bone Age, Deep Neural Networks, Inception-ResNet-V2, Xception Network, Region-of-Interest.

Article History

Received: July 26, 2020

Revised: January 11, 2021

Accepted: January 13, 2021

1. INTRODUCTION

The determination of bone age provides information about an individual's structural and biological maturity. It can be used as a tool for clinical diagnosis of diseases associated with abnormally short or tall stature in children [1] or for forensic purposes. It can also prove to be useful in ascertaining the chronological age if accurate birth records are unavailable. Many deep learning applications have been successful in substituting the former methods.

Traditionally, the Tanner Whitehouse [2] and the Geurich and Pyle [3] methods are widely practiced in clinical assessment and diagnostics; however, these are labor-intensive and time consuming, vulnerable to observer's mishandling. Predictive analysis is carried out on four major ossification regions in hand, namely epiphyses bone, medial carpal, radius, and the ulna. The first three regions drastically vary according to age, sex and ethnicity [4, 5]. The phalangeal analysis is the most suitable in children (above age 6 in females and above age 8 in males) and therefore, computer-aided medical diagnostic (CAD) systems [6 - 8] method can be deemed the best if applied. The associated techniques can pick out relevant aspects from the phalangeal region using a digital hand atlas. The same cannot be applied to children below the ages of 5-7 years since the presence of soft tissue makes the process of segmentation between epiphysis and metaphysis regions [9] difficult. Among the other alternatives that have been explored

is the CAD-based feature extraction from carpal region-of-interest (ROI) of prepubescent children and the related studies have also been positively assessed [10, 11]. However, due to the complexities surrounding limitations of the algorithm, carpal ROI has not yet been incorporated into the bone age assessment process. An interesting study for reconstruction in the field of surgical procedures was carried out by Solari *et al.* [12] which involves reducing postoperative CSF leak.

Deep Learning [13] and its derivatives have been successful in computer vision tasks such as object detection, classification and segmentation [14, 15]. Some valuable articles [16 - 18] have featured efficient means and methods for biomedical image analysis. Deep CNNs comprise pooling and convolution layers that learn hierarchical feature representations from images, followed by an ensemble of fully connected layers and dense layers that are trained on features extracted from previous layers. It has been possible to create innovative algorithms due to the availability of large datasets, most of which consist of detailed annotated features, and these algorithms/methods have increasingly boosted performances of analytical methods. Similar approaches have also been implemented in bone age assessment tasks [19 - 22], including bone segmentation for advanced feature extraction and thereby facilitating better result achievement while leaving negligible error margin rates.

In this work, two different DNN based frameworks for bone maturity estimation on the RSNA dataset constituting of 10,000 X-Ray images of the human hand are evaluated. The process involves a comparative analysis between two networks,

¹ Address correspondence to this author at Department of Electronics Engineering, KIIT University, Bhubaneswar, India; Tel: 7303409222; E-mail: suprava_patnaik@yahoo.com

with the base models as Inception ResNet v2 and Xception pre-trained networks. The methods suggest the superior performance of the Xception model over the Inception model, however, the Inception ResNet v2 model had a better performance during model training. The Mean Absolute Error (MAE) evaluated on the test set with the Xception model achieves best results with a deviation of around 12.583 months, whereas the Inception ResNet v2 results in a test set MAE of around 13.299 months, making the overall procedure more optimized and can thus assist in improved clinical diagnostic evaluations.

2. RELATED WORKS

The standard bone age estimation paradigm is centered around the Geurich and Pyle [3] and Tanner Whitehouse [2] methods [23]. Deep Convolutional Neural Networks [24] have been widely successful in research related to medical imaging. Pan *et al.* [25] applied deep transfer learning techniques such as multi-characteristic CNNs and an ensemble approach on the RSNA dataset for BAA. Their model achieved an MAE of 8.59, 6.96 and 7.35 months on all, male, and female cohorts respectively. Mansourvar *et al.* [26] designed an automated BAA system that used CBIR (Content Based Image Retrieval) and returned an average error rate of -0.170625 years. Rucci *et al.* [27] developed a scheme for bone classification using neural networks in the Tanner Whitehouse method (TW2) [2] but their results were relatively un-satisfactory with an error rate of 1.4 years. Wu *et al.* [28] incorporated two subnets in their deep learning based pipeline on the RSNA dataset: MASK R-CNN for eliminating background noise and a residual attention subnet based on the aforementioned subnet for generating the final predictive output and related visualizations. These techniques, however, are not well-suited for images with low resolution since they do not perform precision-based image segmentation. In a more advanced approach, Thoderg *et al.* [29] proposed the BoneXpert which used a repository of 3000 carefully annotated bone images and on the basis of a combination of shape, intensity and textural features, efficiently determined bone maturity. Pietka *et al.* [30] developed a bone age estimation method using a digital hand atlas. The preprocessing phase yielded epiphyseal/metaphyseal regions of interest (EMROIs) which there then fed to feature extraction functions. Three ratios of distance were generated: ed/md, ed/dist, and md/dist and the final assessment gave near accurate results, with only a detection failure in 4% of the radiographs. Several other such systems/methods have also been designed [31 - 33]. Certain algorithms [34 - 37] have also been established that can be applied in hand-wrist analysis, dealing with segmenting out only certain zones of the radiology images.

DCNN can be efficiently reinforced in tasks related to bone age estimation [38 - 40, 19]. Though some of these techniques give satisfactory results, most of them generally tend to be inclined towards some common shortcomings:

- (1) The techniques might generate bias since the evaluation

is centered around coarse digital processed images of hands bones.

- (2) Most use regressors that are more suitable for low resolution images rather than high quality latent counterparts. This can limit the overall performance of the BAA system.

3. DEEP NEURAL NETWORKS

Multiple assessments suggest incorporation of Deep Neural Network architectures instead of Convolutional Neural Networks. Many researchers in their previous contributions have tried to use Convolutional Neural Networks for the identification of skeletal age from X-Ray images, but the methods involved using space invariant ANN's, based on their shared-weights architecture and translation invariance characteristics. Deep Neural Networks subject to methods involving the transfer of feature maps, layer by layer as supplementary information, to perform batch-wise model preparation. Pre-trained DNN models like Inception-ResNet V2 and Xception are selected as base models, and more convolutional blocks are added to these base models to evaluate them independently.

4. METHODS

This paper proposes a method to identify the age of subjects from hand X-Ray images. This involves incorporation of a comparison analysis of two pre-trained Deep Neural Network classifiers, namely Inception-ResNet V2 and Xception. Different evaluation parameters, such as Mean absolute error (MAE), Mean squared error (MSE), Root mean squared error (RMSE) and R-squared are used to identify the range from their predicted age and a ground truth labelled by the medical experts. The proposed method suggests the performance of models similar to medical experts and are aimed as highly useful tools for computer-aided diagnosis, towards easier age identification. The experiments were conducted using NVIDIA P100 Graphics Processing Unit (GPU). The setup further aids us to conduct experiments faster and gather results at a rapid rate than usual (Fig. 1).

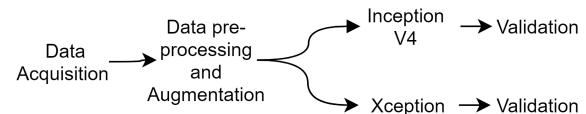


Fig. (1). Flowchart for the proposed model.

4.1. Data description

The RSNA X-Ray data has been collected from Pediatric Bone Age Challenge 2017 competition. The dataset [41] is originally contributed by Stanford University, The University of Colorado and The University of California, Los Angeles. We have taken advantage of a dataset which consists of 12,611 X-Ray images of human hand. The dataset contains hand images for image accession. The sample of X-Rays is shown in Fig. (2).

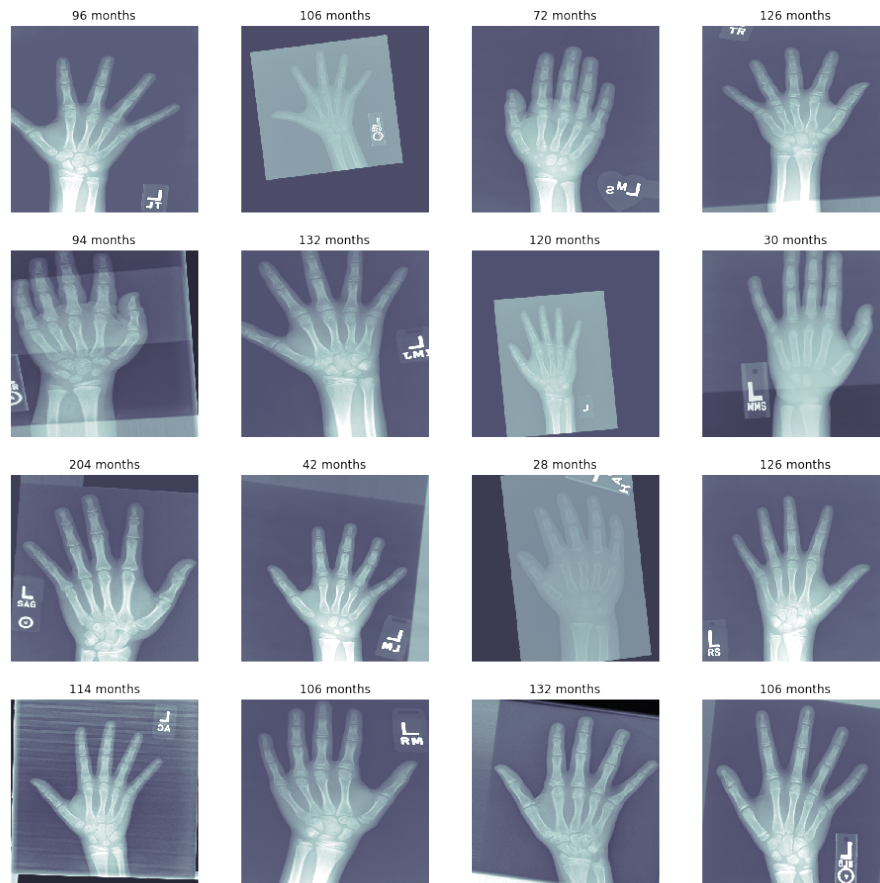


Fig. (2). Dataset description.

4.2. Data Pre-Processing and Augmentation

The X-Ray images were already in high-dimensional format; hence enhancing or distorting features in images was not required. However, each image was resized to 512 X 512 pixels. The images were changed to gray-scale so that the number of channels is reduced to 1, thus affecting complexity of the architectures. Image Data Generator is used to create batches of digital image data by using real-time data. This involves the augmentation of data. The data augmentation strategy increases data diversity for a model to increase their training capacity without any increase in training instances. This method is carried out using enhancement tools like cropping, flip, padding, resizing or changing rotation angle in order to manipulate source data.

The size of the data was reduced from 12,611 images to 10,000 images, which involves the removal of labels having duplicate or erroneous indexes. The new total is split into a training set (6,000 images), a test set (2,000 images), and a validation set (2,000 images). Upon completion of the preprocessing steps up to a sufficient standard, the DNN architectures were finally applied upon the obtained images as part of the main evaluation.

4.3. Model Architectures

The Deep Neural Network architectures used for assessment were chosen because of their optimized performance when compared to other contemporary pre-trained

DNN classifiers. Both the models were initially set to train for 15 epochs. A greater number of epochs was not used since the same can prompt model overfitting, while less number of epochs can bring about an underfit model. This technique permits users to determine huge number of training epochs, training halts and determine when the model shows promising improvement across the validation dataset. Three callback techniques are utilized for model compiling, specifically ModelCheckpoint, EarlyStopping and ReduceLROnPlateau. The ModelCheckpoint callback class allows to define the location and settings to save improved model weights. The EarlyStopping callback is configured when instantiated via arguments. ReduceLROnPlateau callback monitors a parameter and if no improvement is observed for a certain number of 'patients' per epoch, the learning rate is diminished.

4.3.1. Inception ResNet v2

Inception ResNet V2 (Inception ResNet v2) [42] is a deeply convoluted neural network that is a hybrid of Inception and the ResNet modules. Here residual connections are introduced to add the output of the convolutional operations of the inception modules to the input and further the 1 X 1 convolutions are applied after the original convolutions to resemble the depth size. The residual connections replace pooling operations. The stability in the network is maintained by scaling the residual activation functions by values around 0.1 to 0.3.

As the model architecture is deep enough to overfit the data, we have employed dropout layers. These layers randomly drop out some of the nodes to bypass the complexities in the model, which in turn affects our model by losing vital patterns in the data. This problem is evaded by a layer of batch normalization. It normalizes the data to a definite range to dodge covariance shifting. The total number of parameters in the model is reduced by the global average pooling layer, thus decreasing any further chances of overfitting. The loss function is optimized by Adam optimizer [43]. The underlying equations for effective convergence and weight updates using Adam optimizer are explained in Equations 1-4.

Initial weights:

$$\rho_m \leftarrow 1, \rho_v \leftarrow 1, \mathbf{m} \leftarrow \mathbf{0}, \mathbf{v} \leftarrow \mathbf{0} \tag{1}$$

Adam optimiser update equations:

$$\mathbf{m} \leftarrow \beta_m \mathbf{m} + (1 - \beta_m) \nabla_w J \tag{2}$$

$$\mathbf{v} \leftarrow \beta_v \mathbf{v} + (1 - \beta_v) (\nabla_w J \otimes \nabla_w J) \tag{3}$$

$$\mathbf{w} \leftarrow \mathbf{w} - \alpha \left(\frac{\mathbf{m}}{\sqrt{\mathbf{v}} + \epsilon} \frac{\sqrt{1 - \rho_v}}{1 - \rho_m} \right) \tag{4}$$

Here, \otimes refers to element-multiplication and in Equation 4, the operations under the root are also handled element-wise.

This model is trained upon 375 images per batch through 16 such batches during the training phase and verified on 125 images per batch for 16 batches through the validation phase. This batch size is maintained for generalization of the results. This generalization of results helps our model to predict further instances outside the training set.

4.3.2. Xception

The Xception network [44] is a deeply convoluted neural network that uses feature extraction to learn further distinct patterns in the data with a lesser estimate of parameters. The primary principle of Xception network is that it uses cross channel correlation and spatial correlation in a decoupled manner. This architecture is mainly centered around depth

separable convolution accompanied by point-wise convolution, consisting of 36 convolutional layers structured into 14 modules for core feature extraction. The Xception network is the most improved form of the Inception network.

Additionally, batch normalization is used to normalize the input data, which controls the co-variance shift in the specified image data. It also enables the data to learn by itself independently. Batch normalization decreases overfitting by adding some noise in the data which enables us to use lesser dropout values. This saves the data from losing crucial visual patterns in the data. A dropout layer of 0.5 is added after the batch normalization layer. This layer is used to avoid overfitting in the data alongside keeping the crucial information in the data by batch normalization. These layers are followed by a global average pooling layer to avoid overfitting by minimizing parameters as it decreases the overall spatial dimensions of the images, reducing the model complexity for better performance. This layer is succeeded by a fully connected layer of the linear activation function to find the mean absolute error. Here Adam optimizer [43] is employed to learn and reach the global minima for optimizing the loss function. This model uses a batch size of 8, which enables us to use 750 images per training batch and 250 images per validation batch. Sequentially running these mini-batches helps in accumulating variables and updating them in succeeding batches. This helps in optimizing the memory usage and in generalizing the results by detouring from getting stuck in the local minima.

4.4. Evaluation

Both of the DNN models were initially set to train for 15 epochs. However, due to callbacks parameter, the Xception model was trained for 15 epochs, while Inception ResNet v2 was trained for 10 epochs and then, the training was halted. The validation-loss parameter was monitored, and mean absolute error of months was assessed for both training and validation batches per epoch. The evaluation graphs for the performance assessment of both models are demonstrated in Fig. (3).

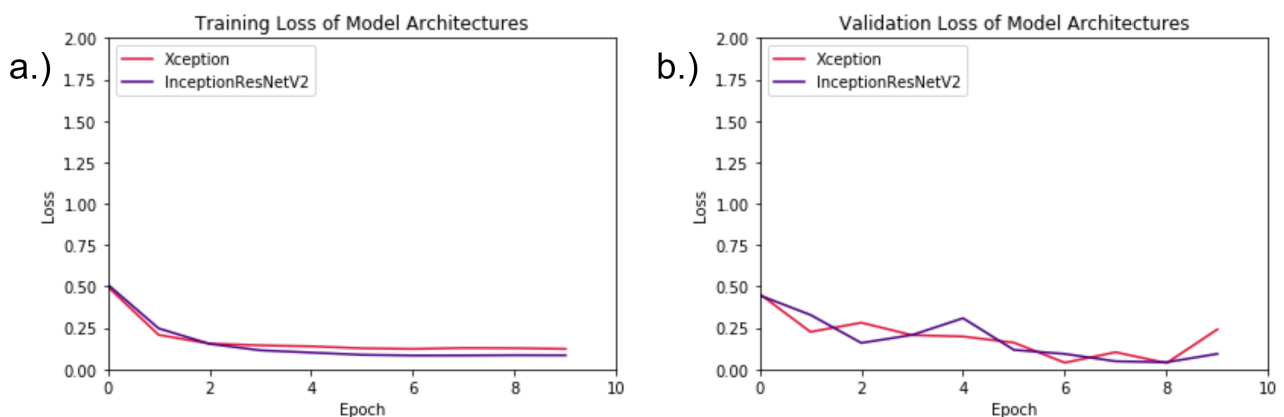


Fig. (3). Evaluating and monitoring individual performance of DNN models vs epochs with respect to a) Training loss, and b.) Validation loss.

5. RESULTS

The DNN models were compared using evaluation parameters, such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R Square value. These four error metrics are used to determine the model with more optimised results.

The Mean Absolute Error (MAE) is a metric used to determine the similarity between two sets.

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - x_i)^2 \tag{5}$$

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |y_i - x_i| \tag{6}$$

The Root Mean Squared Error (RMSE) is a metric for determining the similarity between the true and predicted values. This parameter is similar to MAE, except for two conditions. Firstly, each absolute error is squared before being summed. Secondly, the final result (MSE) is square-rooted before being returned.

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - x_i)^2} \tag{7}$$

The idea of R squared is that if more samples are added, the coefficient will show the probability of a new point falling

on the line.

$$R - Squared = 1 - \frac{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - \bar{y})^2} \tag{8}$$

As shown in Table 1, the parameters suggest almost similar performance of both the DNN models. The performances eventually can be boosted if more data is added to the training set. Xception has a MAE of 12.583 while Inception ResNet v2 has an MAE of 13.299. The final R- Squared value is 0.943 for Xception, whereas Inception ResNet v2 achieves a value of 0.935. All these observations clearly show that Xception has a better chance of age identification than Inception ResNet v2. Fig. (4) shows the distribution of predicted results with respect to ground truth for both models.

The models employed in this system were used to predict the age of the hand X-Rays in terms of months, with pre-determined ground truth set by the medical experts. The results, as demonstrated in Fig. (5), show a deviation from actual age in months in terms of MAE from the ground truth. Both the models present almost similar performances, thus guaranteeing enhanced identification examples with fewer training instances.

Additionally, both models showcase improved performances in terms of R square parameters, which denote the condition that the models already have a higher chance of correctly identifying the age if exposed to and tested upon unseen data.

Table 1. Performance analysis using evaluation metrics.

Metrics Used	Inception ResNet V2	Xception
MSE	287.328	254.025
MAE	13.299	12.583
RMSE	16.951	15.938
R squared	0.935	0.943

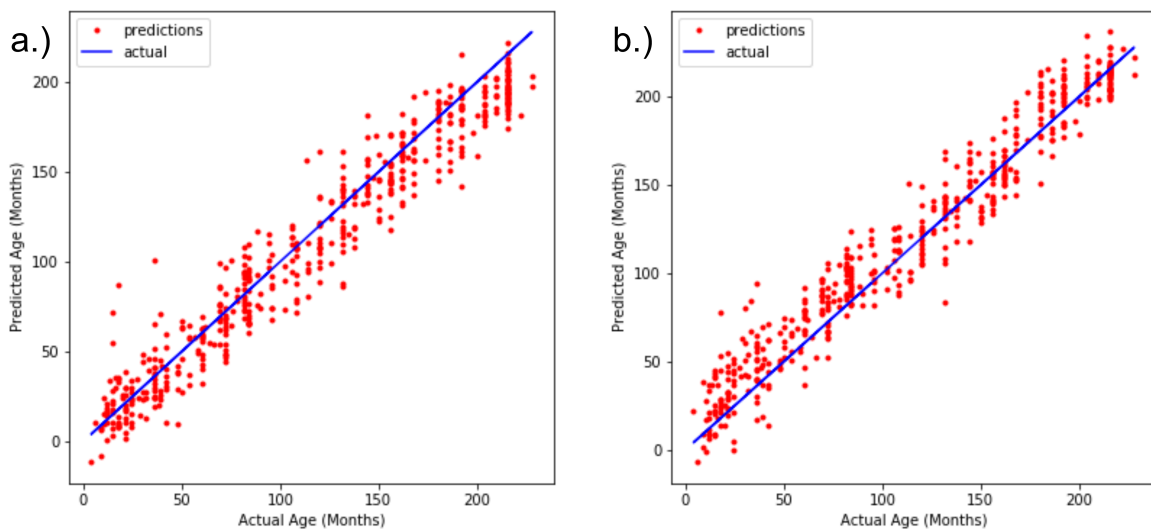


Fig. (4). Predicted results vs actual results of a) Inception ResNet v2, and b) Xception model.

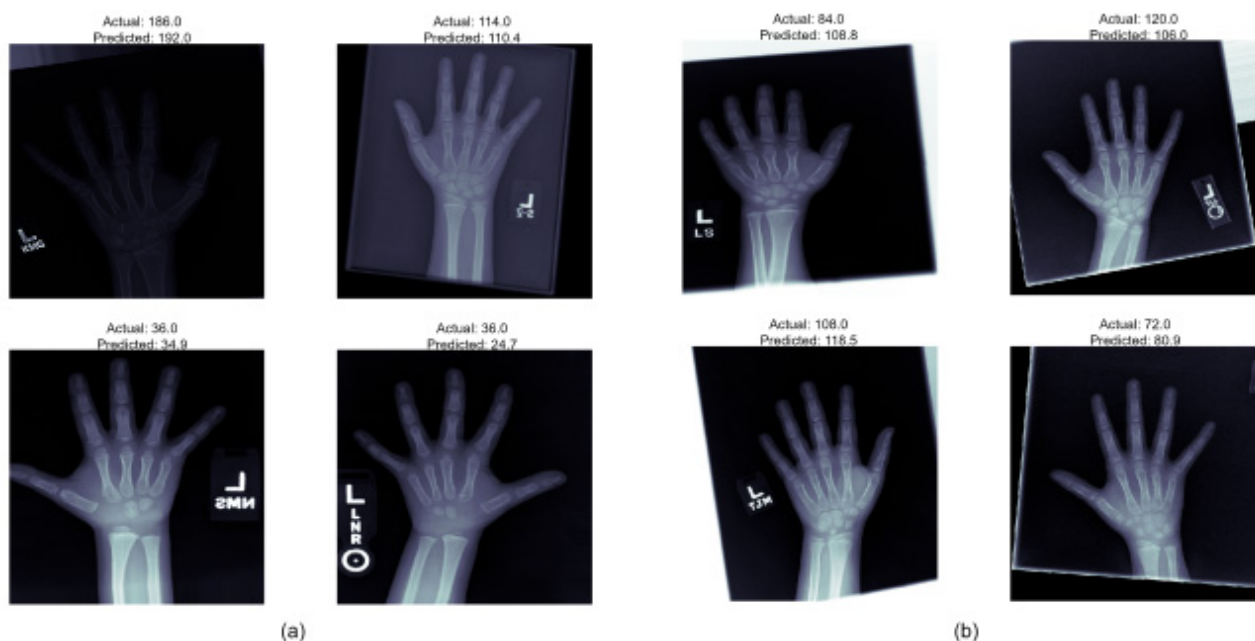


Fig. (5). Predicted age vs actual age of a) Inception ResNet v2, and b) Xception model.

CONCLUSION

We have proposed a fully automated system for efficiently determining skeletal maturity using the RSNA dataset. The system, consisting of two primary models, Xception and Inception ResNet V2, automatically extract relevant features from the data and achieves excellent outcomes in terms of mean absolute error of 12.583 and 13.299, respectively, in the models. The results could be further enhanced if tried with system specifications allowing more reduction of the Learning Rate initially unaffected because of callbacks. The proposed model must also be exposed to more diverse training data to permit model diversity and generalization of the results, which would provide an advantageous assessment of images to make image identifications with reduced mean absolute error.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] D.D. Martin, J.M. Wit, Z. Hochberg, L. Säwendahl, R.R. van Rijn, O. Fricke, N. Cameron, J. Caliebe, T. Hertel, D. Kiepe, K. Albertsson-Wikland, H.H. Thodberg, G. Binder, and M.B. Ranke, "The use of bone age in clinical practice - part 1", *Horm. Res. Paediatr.*, vol. 76, no. 1, pp. 1-9, 2011. [http://dx.doi.org/10.1159/000329372] [PMID: 21691054]
- [2] J.M. Tanner, R. Whitehouse, N. Cameron, W. Marshall, M. Healy, and H. Goldstein, *Assessment of skeletal maturity and prediction of adult height (TW2 method)*, Saunders London, 2001.
- [3] W. Hamilton, "Radiographic atlas of skeletal development of the hand and wrist", *J. Anat.*, vol. 85, no. 1, pp. 85-1951, 1951.
- [4] W.W. Greulich, and S.I. Pyle, *Radiographic atlas of skeletal development of the hand and wrist*, Stanford university press, 1959. [http://dx.doi.org/10.1097/00000441-195909000-00030]
- [5] F.E. Johnston, and S.B. Jahina, "The contribution of the carpal bones to the assessment of skeletal age", *Am. J. Phys. Anthropol.*, vol. 23, no. 4, pp. 349-354, 1965. [http://dx.doi.org/10.1002/ajpa.1330230408] [PMID: 4287192]
- [6] F. Cao, H.K. Huang, E. Pietka, and V. Gilsanz, "Digital hand atlas and web-based bone age assessment: System design and implementation", *Comput. Med. Imaging Graph.*, vol. 24, no. 5, pp. 297-307, 2000. [http://dx.doi.org/10.1016/S0895-6111(00)00026-4] [PMID: 10940607]
- [7] A. Zhang, F. Cao, E. Pietka, B.J. Liu, and H. Huang, "Data mining for average images in a digital hand atlas", In: *Medical Imaging 2004:*

- PACS and Imaging Informatic*, vol. 5371. 2004, pp. 251-258. International Society for Optics and Photonics, 2004. [http://dx.doi.org/10.1117/12.535569]
- [8] H.K. Huang, A. Zhang, B. Liu, Z. Zhou, J. Documet, N. King, and L.W. Chan, "Data grid for large-scale medical image archive and analysis", *Proceedings of the 13th annual ACM international conference on Multimedia*, 2005pp. 1005-1013 [http://dx.doi.org/10.1145/1101149.1101357]
- [9] A. Zhang, A. Gertych, and B.J. Liu, "Automatic bone age assessment for young children from newborn to 7-year-old using carpal bones", *Comput. Med. Imaging Graph.*, vol. 31, no. 4-5, pp. 299-310, 2007. [http://dx.doi.org/10.1016/j.compmedimag.2007.02.008] [PMID: 17369018]
- [10] E. Pietka, L. Kaabi, M.L. Kuo, and H.K. Huang, "Feature extraction in carpal-bone analysis", *IEEE Trans. Med. Imaging*, vol. 12, no. 1, pp. 44-49, 1993. [http://dx.doi.org/10.1109/42.222665] [PMID: 18218390]
- [11] D.R. Kirks, and N.T. Griscom, *Practical pediatric imaging: diagnostic radiology of infants and children.*, Lippincott Williams & Wilkins, 1997.
- [12] D. Solari, L.M. Cavallo, P. Cappabianca, I. Onofrio, I. Papallo, A. Brunetti, L. Ugga, R. Cuocolo, A. Gloria, and G. Improta, "Skull base reconstruction after endoscopic endonasal surgery: new strategies for raising the dam", *2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0&IoT)*, 2019pp. 28-32 [http://dx.doi.org/10.1109/METRO4.2019.8792878]
- [13] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, vol. 521. 2015, no. 7553, pp. 436-444.
- [14] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks", *Advances in neural information processing systems*
- [15] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *arXiv preprint arXiv*, 2014.1409.1556
- [16] S. Koitka, M.S. Kim, M. Qu, A. Fischer, C.M. Friedrich, and F. Nensa, "Mimicking the radiologists' workflow: Estimating pediatric hand bone age with stacked deep neural networks", *Med. Image Anal.*, vol. 64, 2020.101743 [http://dx.doi.org/10.1016/j.media.2020.101743] [PMID: 32540698]
- [17] E. S. of Radiology, "What the radiologist should know about artificial intelligence—an esr white paper", *Insights into imaging*, vol. 10, no. 1, p. 44, 2019.
- [18] L.M. Prevedello, S.S. Halabi, G. Shih, C.C. Wu, M.D. Kohli, F.H. Chokshi, B.J. Erickson, and J. Kalpathy-Cramer, "K. P. Andriole, and A. E. Flanders, "Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions", *Radiology: Artificial Intelligence*, vol. 1, no. 1, p. e180031, 2019.
- [19] V.I. Igloukov, A. Rakhlin, A.A. Kalinin, and A.A. Shvets, Paediatric bone age assessment using deep convolutional neural networks. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, pp. 300-308. [http://dx.doi.org/10.1007/978-3-030-00889-5_34]
- [20] X. Ren, T. Li, X. Yang, S. Wang, S. Ahmad, L. Xiang, S.R. Stone, L. Li, Y. Zhan, D. Shen, and Q. Wang, "Regression convolutional neural network for automated pediatric bone age assessment from hand radiograph", *IEEE J. Biomed. Health Inform.*, vol. 23, no. 5, pp. 2030-2038, 2019. [http://dx.doi.org/10.1109/JBHI.2018.2876916] [PMID: 30346295]
- [21] A. Gertych, A. Zhang, J. Sayre, S. Pospiech-Kurkowska, and H.K. Huang, "Bone age assessment of children using a digital hand atlas", *Comput. Med. Imaging Graph.*, vol. 31, no. 4-5, pp. 322-331, 2007. [http://dx.doi.org/10.1016/j.compmedimag.2007.02.012] [PMID: 17387000]
- [22] S. Mutasa, P.D. Chang, C. Ruzal-Shapiro, and R. Ayyala, "Mabal: a novel deep-learning architecture for machine-assisted bone age labeling", *J. Digit. Imaging*, vol. 31, no. 4, pp. 513-519, 2018. [http://dx.doi.org/10.1007/s10278-018-0053-3] [PMID: 29404850]
- [23] G.R. Milner, R.K. Levick, and R. Kay, "Assessment of bone age: A comparison of the Greulich and Pyle, and the Tanner and Whitehouse methods", *Clin. Radiol.*, vol. 37, no. 2, pp. 119-121, 1986. [http://dx.doi.org/10.1016/S0009-9260(86)80376-2] [PMID: 3698492]
- [24] G. Huang, Z. Liu, L. Van Der Maaten, and K.Q. Weinberger, "Densely connected convolutional networks", *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708. [http://dx.doi.org/10.1109/CVPR.2017.243]
- [25] X. Pan, Y. Zhao, H. Chen, D. Wei, C. Zhao, and Z. Wei, "Fully automated bone age assessment on large-scale hand x-ray dataset", *Int. J. Biomed. Imag.*, . [http://dx.doi.org/10.1155/2020/8460493]
- [26] M. Mansourvar, R.G. Raj, M.A. Ismail, S.A. Kareem, S. Shanmugam, S. Wahid, R. Mahmud, R.H. Abdullah, F.H.F. Nasaruddin, and N. Idris, "Automated web based system for bone age assessment using histogram technique", *Malays. J. Comput. Sci.*, vol. 25, no. 3, pp. 107-121, 2012.
- [27] M. Rucci, G. Coppini, I. Nicoletti, D. Cheli, and G. Valli, "Automatic analysis of hand radiographs for the assessment of skeletal age: A subsymbolic approach", *Comput. Biomed. Res.*, vol. 28, no. 3, pp. 239-256, 1995. [http://dx.doi.org/10.1006/cbmr.1995.1016] [PMID: 7554858]
- [28] E. Wu, B. Kong, X. Wang, J. Bai, Y. Lu, F. Gao, S. Zhang, K. Cao, Q. Song, and S. Lyu, "Residual attention based network for hand bone age assessment", *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019pp. 1158-1161 [http://dx.doi.org/10.1109/ISBI.2019.8759332]
- [29] H.H. Thodberg, S. Kreiborg, A. Juul, and K.D. Pedersen, "The BoneXpert method for automated determination of skeletal maturity", *IEEE Trans. Med. Imaging*, vol. 28, no. 1, pp. 52-66, 2009. [http://dx.doi.org/10.1109/TMI.2008.926067] [PMID: 19116188]
- [30] E. Pietka, A. Gertych, S. Pospiech, F. Cao, H.K. Huang, and V. Gilsanz, "Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction", *IEEE Trans. Med. Imaging*, vol. 20, no. 8, pp. 715-729, 2001. [http://dx.doi.org/10.1109/42.938240] [PMID: 11513023]
- [31] K. Hill, and P.B. Pynsent, "A fully automated bone-ageing system", *Acta Paediatr. Suppl.*, vol. 406, pp. 81-83, 1994. [http://dx.doi.org/10.1111/j.1651-2227.1994.tb13430.x] [PMID: 7734817]
- [32] K. SATO, "Setting up an automated system for evaluation of bone age", *Endocrine j.*, vol. 46, no. Suppl, pp. S97-S100.
- [33] D.J. Michael, and A.C. Nelson, "HANDX: a model-based system for automatic segmentation of bones from digital hand radiographs", *IEEE Trans. Med. Imaging*, vol. 8, no. 1, pp. 64-69, 1989. [http://dx.doi.org/10.1109/42.20363] [PMID: 18230501]
- [34] S.N. Cheng, H-P. Chan, L.T. Niklason, and R.S. Adler, "Automated segmentation of regions of interest on hand radiographs", *Med. Phys.*, vol. 21, no. 8, pp. 1293-1300, 1994. [http://dx.doi.org/10.1118/1.597402] [PMID: 7799874]
- [35] G. Manos, A. Cairns, I.W. Ricketts, and D. Sinclair, "Automatic segmentation of hand-wrist radiographs", *Image Vis. Comput.*, vol. 11, no. 2, pp. 100-111, 1993. [http://dx.doi.org/10.1016/0262-8856(93)90076-S]
- [36] J. Duryea, Y. Jiang, P. Countryman, and H. Genant, "Automated algorithm for the identification of joint space and phalanx margin locations on digitized hand radiographs", *Med. Phys.*, vol. 26, no. 3, pp. 453-461, 1999. [http://dx.doi.org/10.1118/1.598537]
- [37] P. Peloschek, G. Langs, M. Weber, J. Sailer, M. Reisegger, H. Imhof, H. Bischof, and F. Kainberger, "An automatic model-based system for joint space measurements on hand radiographs: Initial experience", *Radiology*, vol. 245, no. 3, pp. 855-862, 2007. [http://dx.doi.org/10.1148/radiol.2452061281] [PMID: 17951347]
- [38] H. Lee, S. Tajmir, J. Lee, M. Zissen, B.A. Yeshiwas, T.K. Alkasab, G. Choy, and S. Do, "Fully automated deep learning system for bone age assessment", *J. Digit. Imaging*, vol. 30, no. 4, pp. 427-441, 2017. [http://dx.doi.org/10.1007/s10278-017-9955-8] [PMID: 28275919]
- [39] D.B. Larson, M.C. Chen, M.P. Lungren, S.S. Halabi, N.V. Stence, and C.P. Langlotz, "Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs", *Radiology*, vol. 287, no. 1, pp. 313-322, 2018. [http://dx.doi.org/10.1148/radiol.2017170236] [PMID: 29095675]
- [40] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in X-ray images", *Med. Image Anal.*, vol. 36, pp. 41-51, 2017. [http://dx.doi.org/10.1016/j.media.2016.10.010] [PMID: 27816861]
- [41] S.S. Halabi, L.M. Prevedello, J. Kalpathy-Cramer, A.B. Mamonov, A. Bilbily, M. Cicero, I. Pan, L.A. Pereira, R.T. Sousa, N. Abdala, F.C. Kitamura, H.H. Thodberg, L. Chen, G. Shih, K. Andriole, M.D. Kohli, B.J. Erickson, and A.E. Flanders, "The rsna pediatric bone age machine learning challenge", *Radiology*, vol. 290, no. 2, pp. 498-503, 2019. [http://dx.doi.org/10.1148/radiol.2018180736] [PMID: 30480490]
- [42] C. Szegedy, S. Ioffe, V. Vanhoucke, and A.A. Alemi, "Inception-v4,

- inception-resnet and the impact of residual connections on learning", *Thirty-first AAAI conference on artificial intelligence*, 2017
- [43] D.P. Kingma, and J. Ba, "Adam: A method for stochastic optimization", *arXiv preprint arXiv*, vol. 1412.6980, 2014.
- [44] F. Chollet, "Xception: Deep learning with depthwise separable convolutions", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017pp. 1251-1258 [http://dx.doi.org/10.1109/CVPR.2017.195]

© 2021 Patnaik. *et al.*

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.