Bone Age Assessment using Deep Learning

[1]Mariyam Sunny Kattukaran, [2]Abey Abraham
[1,2]Dept. of IT, Rajagiri School of Engg. & Technology, APJ Abdul Kalam Technological University, India

Abstract
The assessment of skeletal bone age is very subjective and tiring process. As a result, computer assisted techniques are developed to replace hand operated examination techniques in medical industry. The research aims to find out a new computer aided technique which is based on convolutional neural network. It is developed to check the skeletal maturity as well as gender and race. The most important advantage of this proposal is that it minimizes the segmentation problem bared by the existing systems. Skeletal bone age assessment is the most used clinical practice to scrutinize endocrinology, genetic and growth disorders in the youngsters. It is usually executed using radiological analysis of left hand wrist using Greulich and Pyle technique or Tanner Whitehouse technique. Nevertheless, both the techniques have many cons including human observation variability and consumption of time.

The study proposes deep learning techniques to check skeletal bone age and also to generalize the gender and race. This bone age assessment technique works on public data set for all age ranges, races and genders.

Keywords
Convolutional Neural Networks, Deep Learning for Medical Images, Tanner White House,Greulich and Pyle

I. Introduction

Bone age assessment or bone maturity analysis is a scientific application that can be used to check the skeletal development in youngsters. As a result of the non-efficiency to analyse the maturity age of the children using actual age, the maturation of the skeletons is used as a signage for growth problems. The X ray image of the left hand wrist is used because it is a trustable index of skeletal maturation and nevertheless can be used to indicate the biological development of the bones based on ossification regions and calcium deposits in the ossification area. The endocrine disorders, chromosomal disorders and early sexual maturity can be identified using the differences between the calculated skeletal bone age and biological age in children.

BAA is a radiological process to check the ossification maturity in the left hand using X ray and then by identifying the bone age with the help of an Atlas that contains hundreds of standard images. The differences in the bone age and chronological age is the indicator of growth disorders, chromosomal disorders, endocrine disorders that could be identified in the early stages with the help of BAA techniques. Even though bone age assessment is an important routine in clinical application it has not been have improved for more than thirty five years. The Greulish-Pyle (GP) and Tanner Whitehouse (TW2) methods are the much known manual methods used in the early period.

In GP method the patient's bone radiographs are compared with standard radiographs with the help of an Atlas and make final decisions, while in TW2 method, we use scoring process. The main limitation of both the methods are inter and intra observation variability as well as it is much time consuming. Normally a radiologist performs bone age evaluation to check child's maturity cannot be assured about the accuracy of the analysis. As a result, this is the most important factor for the presentation of computer assisted method to estimate the maturity of the bones. Nonetheless, the automated bone age system is still inside the empirical time caused of non-adequate performance of the system. Some of the recommended techniques have been analyzed in the literature review.

The discrepancies between the skeletal bone age and biological age are identified using the radiological examinations which are called bone age assessment systems. The differences depict the deformities in the skeletal growth of the youngsters or the hormonal disorders. To calculate bone age, on the basis of bone age and the ability of reproduction is a difficult process and consumes more time in clinical procedures. The bone age assessment considers three vital factors. (a) Presence of primary and secondary centers of ossification (b) development of both centers (c) timing of the union of primary and secondary centers.

The presence and the variations in the above factors are clearly found in the dry bone and radiographic images. The verdict about the accuracy of the age can be assured on the basis of development timing condition of the ossification region and the epiphycal union identification which is dependent on whether the dry bone is being monitored or can be visualized by an imaging method such as radiograph or X-ray. The analysis of chronological age is matched, which contain a comparison of a radiograph image with reference to a defined sample of known age and sex. The act of age calculation is basically used to indicate the biological maturity which is then changed to chronological age with the help of reference that is used for comparison processes.

The data used for the age calculation are collected from different sources are presented as a sequence called an "Atlas". Most of the processed information are used to formulate the atlas were compiled at the time of longitudinal studies during the 1900's. These information was collected for each individual youngsters as part of an anthropometric exercise in the structure of uniform radiograph. Therefore, the object was to portray the growth of normal life and all the candidates had a history of health without any abnormalities or disease. The compiled data provides reference data to calculate the age of an unknown youngster for identification or research purposes.

Factors like the environment and the nutritional conditions which vigorously affect the progress of the children in the community. The atlases which are created on the basis of the health of the adolescents who had healthy food to consume were considered and were ideal to be used as the benchmark for comparison. These atlases contained a collection of images which showed the maturity differences of the bones as a strong motivator for the calculation of the age. The atlases considered the step of maturation of the child with the unidentified age and it helps to identify the most ideal age of a child with given age. This is a usual practice for almost all the atlases and the age found out by the experts is based on it accordingly. The meaning of using the atlas means that they show a non-permanent depiction of the maturity tempo of the youngsters of the known race. The difficulty is whether this processed data is important to the modern community or it can be used for children of various races with a different regime and medical treatments.

State of the art review portray that the atlases for finding out the age is based on the radiograph image of the wrist of the left hand. These kinds of studies can be divided on the basis of techniques to check the accuracy the classification technique contains the below activities: (a) Checking age analysis techniques on a special community. (b) Comparison of the precision of various atlases from the similar skeletal area on the similar group. (c) Comparison of error observer.(d) comparison of the maturation levels on the various parts of the body in the similar group.

Albeit, analysis of bone age is flexible from a number of bones in the body like the elbow, pelvic, clavicle, foot, shoulder, or ankle, nevertheless the high expenses and time consumption and the risk of exposure to the radiation depicts that this is not ideal or pragmatic for the former researchers to consider for BAA.

III. Related Work

As aforementioned, bone age is defined as an indicator of skeletal maturity using radiography of the ossification center. Despite a large volume of scientific research on BAA, there is a lack of agreement concerning the accuracy of bone age methods which is acceptable for a clinical environment . For BAA in both clinical environments and courts of justice, it is important to yield the most accurate result. An automated bone age system could reasonably eliminate the role of a human observer, which would decrease the subjectivity in assessment as the main reason for the loss of accuracy . This part of the discussion classifies the computerized methods for BAA, which is the significant topic of this research.

Most of the automated systems for estimation of bone age derived the state of skeletal maturity from X-ray images of the left hand wrist. This is not an easy task because the hand wrist includes a group of various bones, which rapidly change shape over time, and also some bones overlap with maturation. As mentioned previously, analysing bone age is a complicated process even for experts . Most computer-based methods use the TW approach due to the scoring for skeletal maturity and separate stages. Specific image processing techniques are needed to assess the radiograph of a known hand .

Researchers admit the significance in automating the method for the estimation of bone age. These methods use some intelligent techniques, such as segmentation of the hand, while some are only used in the research environment. It is estimated that computerized methods in BAA could decrease the cost of assessment of bone age through a decrease in the time that radiologists spend in predicting the bone age.

A. HANDX System

This is the first semiautomated system for BAA was introduced by Micheal and Nelson in 1989. The authors claim that this system, which they call HANDX, is able to automatically segment bones in X-ray images of the hand wrist using image processing techniques. This system reduces the variability of the observer and the output is useful to detect abnormalities of skeletal growth in children. This computer vision system works in three parts: preprocessing, segmentation, and measurement. In the first stage the radiographs are normalized to feed in the second step. The segmentation stage finds the specific bones in the hand and also isolates the edges of the bone, and, finally, quantitative parameters are achieved in the last stage. This semiautomated system has no reasonable accuracy when the hand image is fused and has never been evaluated on a large scale.

ADVANTAGE: Reduced observation variability

DISADVANTAGE:No reasonable accuracy

B. PROI-Based System

PROI is the region that includes the phalanges and epiphyses. For the estimation of bone age, in the first process the system scans a horizontal line and the lower boundary of the PROI is found before the soft tissue between the thumb and first finger is detected. In the next stage, the upper boundary containing a horizontal line at the edge of the third finger is scanned. When the upper, lower, left, and right boundary of the PROID have been detected, the segmentation stage starts. A gradient image is used for segmentation of the bones and the output threshold is based on empirical analysis to determine the bone edges. The density of value of pixels at the end of the region is higher than the center section. In this method, the boundary between the third distal, middle, and proximal phalanges is measured. This measurement uses the standard table prepared by the Garn group involving phalangeal length to convert into skeletal age. The system has been evaluated by 50 computer radiographs (CR) of patients and a comparison of the results with an observer (radiologist). The mean difference yielded from the evaluation was 0.02 mm with a measurement error of 0.08 mm .

ADVANTAGE: Low mean difference and error rate.

DISADVANTAGE: Evaluated in small scale.

C. The CASAS System

In 1994, Tanner and Gibbons proposed a computer-based skeletal age scoring system (CASAS), based on the Tanner and Whitehouse2 (TW2) method using radius, ulna, and short bones (RUS) . This semi-automated system digitized X-ray images with a light box and monochrome video camera. Every bone is located on the digital camera using an overlay pattern. The computer assesses the bone age by matching and finding the best average based on fast Fourier transform. The result minimizes the root-mean-square error between the coefficients of the Fourier transform from the unknown bone and Fourier transform of the available bone templates. The patterns are produced by averaging the Fourier transform coefficients using 10 images from bone stages. The system improves to five-root-mean square by using a Gaussian function. The images are only applied to develop a standard skeletal maturity for TW and not for developing the actual bone scoring style. However, in the system the templates have a vital role and selection of the source for making the template is very important. The CASAS system has been tested and evaluated using X-ray images from children in normal and a stable pathologic position. There has been some research on a comparison between the CASAS method and the manual TW method and the results present a reasonable assertion that the CASAS estimation is more accurate than the manual TW method. The conclusion about the CASAS system is that it presents a suitable method for BAA for children in a normal situation. The system is based on a very simple image processing process and the method allows repeatability. The most important weakness is that the method does not work for assessing pathological problems due to deformation of the bone. The method also decreases the assessment objectivity because of the huge number of manual interventions.

ADVANTAGE: More accurate than manual TW method.

DISADVANTAGE: Did not work for assessing with pathological problem.

D. Neural Network Based on the Radius and Ulna

This is a computer-based ystem to predict the bone age based on the TW method and using the radius and ulna. This system is assisted with manual landmarks and then applies an adaptive clustering technique for segmentation of the radius and ulna. The method applies neural networks in the decision state to make a posteriori probabilities that predict the error rate; this feature is specific to this method. The range of the mean difference of the system and observers is large and this method is limited to just four TW3 levels. However, the researchers claim that their method could be extended by improving the bone segmentation. This method proposes that a neural network is valuable for further investigation.

ADVANTAGE:Improving the bone segmentation.

DISADVANTAGE:Limited to four TW3 levels.

E. BoneXpert System

The BoneXpert system is another automated method for BAA that was proposed in 2009. This method works based on shape-driven active appearance and the TW RUS-based approach (using the radius, ulna, and short bones) . The shape and intensity features make a robust algorithm of the active appearance model. A set of components of more than 3,000 bone contours are rotated and scaled, based on the Gobar filters which the parameters are formed in the active appearance model. Thirty coefficients were chosen for features of images using a linear regression technique fed into the active appearance model. Although the usability of the system is still under evaluation, preliminary testing shows that the performance is reasonable and that the accuracy is stated as 0.42 years for using the Greulich & Pyle (GP) method and 0.80 years for using the TW2 method. The rejection rate of the system was about 1% for poor quality but it increased to 18% in some cases for the radius and ulna. The specific point of this method is that it assesses the accuracy of the bone age utilizing the relationship between the X-ray image and linear growth. The standard deviation calculated was 0.5 years, which showed a jump in reproducibility calculated by automated method.

ADVANTAGE: High accuracy.

DISADVANTAGE: Rejects images.

F. WebBased System Using Histogram

This is a fully automated BAA system that uses compression techniques based on the histogram techniques. This approach works on an image repository and similarity measures and uses a content-based image retrieval (CBIR) method for image processing. The system includes a knowledge base consisting of 1100 hand X-ray radiographs classified by gender as well as ethnicity. This approach overcomes the segmentation problem by utilizing a histogram . The evaluation presented 0.170625 years for error rate of the system thereby indicating that this method is a credible method for BAA. However, the system is not reliable for images with poor image quality or abnormal bone structure.

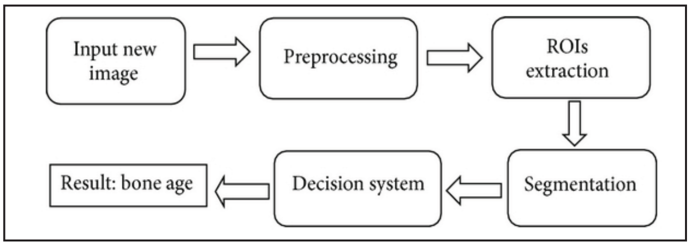

Fig. 1: General Model of BAA Systems

IV. Proposed Work

The model consists of a CNN trained from scratch on the X-ray scan dataset. The advantage of training a new CNN over finetuning an existing one lies in the possibility to tweak the network architecture to the type of images at hand: optimizing the network for grayscale images, 240 reducing the number of layers, letting the network learn specific filters instead of adapting more generic ones. We tested several network architecture and then chose the one which achieved the best accuracy as the final model for evaluation . As convolutional network, we employed a sequence of the following modules: – A pre-trained convolutional layer obtained as a grayscale version. Although CNNs are more commonly used in classification tasks, bone age assessment is a regression task by nature. In order to access performance in both settings, we compare two types of CNNs: regression and classification. Both models share similar architectures and training protocols, and only differ in two final layers.

Our model is a CNN with a regression output. This network represents a stack of six convolutional layers with 32, 64, 128, 128, 256, 256 filters followed by two fully connected layers of 2048 neurons each and a single output . The input size varies depending on the considered region of an image, For better generalization, we apply dropout layers before the fully connected layers. For regression targets, we scale bone age in the range [−1, 1]. The network is trained by minimizing Mean Absolute Error (MAE)

The classification model is similar to the regression one, except for the two final layers. First, we assign each bone age a class.. The second to the last layer is a softmax layer with 240 outputs. This layer outputs vector of probabilities of 240 classes. The probability of a class takes a real value in the range [0, 1]. In the final layer, the softmax layer is multiplied by a vector of distinct bone ages uniformly distributed over 240 integer values [0, 1, ..., 238, 239]. Thereby, the model outputs single value that corresponds to the

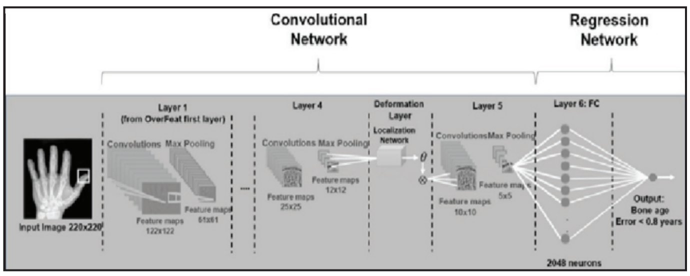

Fig. 2:

V. Result

Fig. 5: Final Result of BAA

In the training the images are loaded and then preprocessing is done.the mosel consist of 6 convolutional layers and one regression network which helps to output single value result from 2048 connected neurons.max-pooling layer and RELU layer is used as part of optimization.

In the testing phase, the performance os the model is evaluated. Fig5 shows the output that comes as result of inputing an X-ray image of eight years old child.the outcome consist of age,gender and race of the respective patient.

VI. Conclusion

In this study we have investigated the application of deep learning to medical images, and in particular for automated skeletal bone age assessment using X-ray images. We have tested convolution neural networks on a dataset of about 1,400 X-ray images and proved that deep learning solutions, even trained on general imagery, are able to cope effectively with all possible cases of automated skeletal bone age assessment . We also designed and trained from scratch several a custom CNN which proved to be the most effective and robust solution in assessing bone age across races, age ranges and gender.in our system ,the segmentation problem is removed by feature mapping technique. In particular, our system consists of six convolution layers, and one regression network. We train this model using the same protocol as the regression model.

Each “convolution layer” includes a ReLU nonlinearity and a max-pooling layer. Different tested network architectures vary by the presence position of the deformation layer, the number of convolution layers, and the number of 275 feature maps. According to the results the architecture of the best performing CNN, consists of six convolution layers, a 2048-neuron fully-connected layer followed by the single-neuron layer providing the estimate fully connected layer followed by a single output neuron.

References

[1] Adeshina, S., Lindner, C., Cootes, T.,"Automatic segmentation of carpal area bones with random forest regression voting for estimating skeletal maturity in infants", In Electronics, Computer and Computation (ICECCO), 2014 11th International Conference on, pp. 1–4, 2014.

[2] Bocchi, L., Ferrara, F., Nicoletti, I., Valli, G.,"An artificial neural network architecture for skeletal age assessment", In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, pp. I–1077–80, Vol. 1, 2003.

[3] Giordano, D., Leonardi, R., Maiorana, F., Scarciofalo, G., Spampinato, C.,"Epiphysis and metaphysis extraction and classification by adaptivthresholding and DoG filtering for automated skeletal bone age analysis", ConfProc IEEE Eng Med Biol Soc, pp. 6552–6557, 2007.

[4] Goodfellow, I. J., Warde-Farley, D., Mirza, M., Courville, A. C., Bengio, Y.,"Maxout networks. In Proceedings of the 30th International Conference on Machine Learning", ICML 2013, Atlanta, GA, USA, 16-21 June 2013, pp. 1319–1327, 2013.

[5] Mansourvar, M., Shamshirband, S., Raj, R. G., Gunalan, R., Mazinani, I.,"An Automated System for Skeletal Maturity Assessment by Extreme Learning Machines", PLoS ONE, 10 , e01384, 2015.

[6] Carty, H.,"Assessment of skeletal maturity and prediction of adult height (tw3 method).: 3rd Edition. Edited by J. M. Tanner, M. J. R. Healy, H. Goldsteinand N. Cameron. pp. 110. London, etc: W. B. Saunders, 2001. ISBN: 0-7020-550 2511-9. 69.95. Journal of Bone and Joint Surgery, British, Vol. 84-B, 310–311.

[7] Van Ginneken, B., Setio, A. A., Jacobs, C., Ciompi, F.,"Off-theshelfconvolutional neural network features for pulmonary nodule detection 570 in computed tomography scans. In Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on pp. 286–289. IEEE, 2015

[8] Ciompi, F., de Hoop, B., van Riel, S. J., Chung, K., Scholten, E. T., Oudkerk,M., de Jong, P. A., Prokop, M., & van Ginneken, B. (2015). Automatic classificationof pulmonary peri-fissural nodules in computed tomography using an ensemble of 2d views and a convolutional neural network out-of-the-box. Medical Image Analysis.

[9] Greulich, W. W., Pyle, S. I.,"Radiographic atlas of skeletal development of the hand and wrist", The American Journal of the Medical Sciences, 1959.

Mariyam Sunny Kattukaran currently M.Tech Network Engineering Student at, Rajagiri school of engineering & technology, Kerala, India. Received B.Tech. degree in Information Technology from Rajagiri School of Engineering & Technology, APJ Abdul kalam technological university, Kochi, Kerala in 2015, Her area of interest includes Deep learning.