e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

HAND BONE AGE ESTIMATION USING ENERGY BASED SEGMENTATION AND MACHINE LEARNING

Nikita Mahajan[*1], Kalyani Mahajan[*2], Manya Nagariya[*3], Kumud Wasnik[*4]

[*1,2,3]B.Tech, SNDT University, CST Department, Usha Mittal Institute Of Technology, Mumbai, Maharashtra, India.

[*4]Professor, SNDT University, CST Department, Usha Mittal Institute Of Technology, Mumbai, Maharashtra, India.

ABSTRACT

In variety of ways Bone age assessment can be useful . It can help pediatricians to identify diseases predict growth, puberty entrance, and assess if a person lacking proper identification is a minor or may be not. To intra -observer variation It is a time-consuming process that is also prone and which can cause problems in many ways. A dataset consisting of 12611 X-ray hand images of persons ranging from new-born to 19 years of age was used. We have used Energy based Segmentation it consists of finding the optimal segmentation according to an energy function, secondly there goes K-means Clustering it is the most commonly used clustering algorithm and for good results CNN and some Python Modules are used .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. Further by setting our epochs to 50 ,and by implementing 50 epochs we got the most precise hand bone age detection.

Keywords: Deep Learning, Object Detection, Bone Age Assessment, Image Processing, CNN.

I. INTRODUCTION

Bones in our body are living tissues ,as well proteins, minerals and vitamins make up the bone .Bone age is the degree of maturation of a child bones .As person grows from childhood ,puberty and finishes growth as a young adult the bones of the skeleton change in size and shape . Bone development is influenced by a number of factors, including nutrition, hormonal secretions, and genetics, Bone age assessment is used to radio logically assess the biological and structural maturity of young patients using the hand as well as wrist x-ray appearances. A valuable tool in assessing children's health in continuing Bone age and also in detailed morphological analyses of left hand x-ray is determined by Bone age using, energy based segmentation, discrete wavelet transformation or image transformation Bone age assessments for Medical Conditions Bone age assessments have various usages in healthcare. It can be used to identify delayed or advanced bone age. Endocrine Disorders are commonly associated with delayed bone age, meaning that a normal bone age can help rule out many such disorders. A tool to rule out certain conditions can be used by bone age assessment. . It can also indicate certain diseases, but most of them demand further investigation and testing for a diagnosis. Children with chronic diseases may have delayed bone age, but not necessarily so. Bone age assessments are also useful when it comes to treating children with short stature. There are five different layers in CNN Input layer : Input layer in CNN should contain image data, and this image is represented by Three dimensional matrix . If we have m training examples then dimension of input will become (54, m). Convolutional layer): It is sometimes called feature extractor layer because features of the image get extracted in this layer. Convolution operation is performed on the image so features like edges can be detected. The output will act as the input for the next layer. Convolutional layer also contains ReLU activation function that will activate neuron based on particular condition Pooling layer: If we apply fully connected layer after Convolutional layer without applying pooling or min/max pooling, then it will be computationally expensive. So, the min/max pooling is only way to reduce the spatial volume of the given input image. Fully connected (FC) layer: which compiles the data which is extracted by previous layers to form the final output. Softmax/logistic layer: Softmax or Logistic layer is basically the last layer of CNN. It resides at the end of FC layer. Output layer: It is the last layer that produces the output label for the input dataset Deep learning is a great way to classify images which is used to build a convolutional neural network (CNN).To build a CNN Python makes it simple by using Keras library ,by summing up the multiplication values and a convolution multiplies a matrix of pixels with a filter matrix or kernel. This process is repeated until all image pixels have been covered and convolution slides over to the next

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

pixel. These algorithms are based on Epiphysis ,Metaphysis that is the ROI (EMORI) Extraction ,ulna and radius out of consideration and leave the carpal ROI , a number of algorithms for automated skeletal bones age exist in the literature Epiphysis is the rounded end of a long bone at joints with adjacent bones ,between the diaphysis and epiphysis the long midsection of the longbones lies the Metaphysis including epiphyseal plate. In all the developed method the algorithms are divided in several step: image pre-processing, background removal, orientation correction, image segmentation and features analysis. Then, in each extracted region a set of isotropic filters based on a Difference of Gaussian algorithm (DoG) with the standard deviation defined by the user are applied, for approximate thresholding of the region, DOG is a feature enhancement algorithm . A small variation of the shape of these bones can change the evaluation of the stage and hence of the final age. However, relying only on the analysis of the EMROI may not be enough for the skeletal bone age evaluation. Sometimes the evaluation of different radiologists can differ. For this reason as a future work it is necessary to implement the automated extraction and classification of the bones. For training a model , we will use the 'fit()' function on our model with the following parameters: training data ,target data , validation data, and the number of epochs. For validation data, test set that is provided is used in the dataset. The number of epochs is the number of times the model will cycle through the data. To improve the model ,the more epochs we run ,and that to a certain point .After that point, the model will stop improving during each epoch.

II. METHODOLOGY

After studying various researches which are made for bone age assessment we decided that a system which is completely automated is required to estimate the bone age. The region of interest which is considered for bone age estimation in this proposed system are phalangeal features, (EMROI) and carpal bones features (CROI). The Energy based segmentation approach is one wherein the energy of reconstructed surface(s)which is obtained after applying the above geometry operations is directly used to segment the data. This algorithm combined with K-means clustering which is another important part of machine learning is put to use in this system. Clustering algorithms are basically unsupervised algorithms but they are similar to Classification algorithms however the basis is different. In Clustering, we are unaware of what we are looking for, and thus we try to identify some segments or clusters in the RSNA data. K-Means clustering algorithm is an unsupervised algorithm popular in machine learning and it is used to segment the area of interest from the background of a given image. The algorithm clusters or partitions the given data into K-clusters or parts which is based on the centroids. The goal here is to find certain groups based on some kind of similarity in the data with the number of groups represented by k. The input images are then compared with standard images and bone age is obtained via result phase and after this the process is repeated for different set of images in order to provide more accuracy. These pre-processed images then serve as an input to CNN (Convolutional Neural Network) and further feature extraction is done by training the data and predicting the hand bone age. As we know CNN is used to extract features from raw images and it requires minimal pre-processing. Also, CNN is able to recognize patterns in the image which are not provided as training data before, provided it resembles one of the training data images. But since we already have pre-processed image using energy based segmentation as mentioned above it will help the CNN to extract features even more precisely which is the ultimate goal of the model.

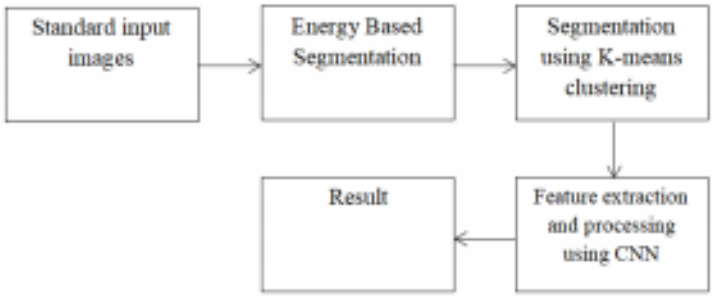

Figure 1: Functional diagram of the model

A. Energy Based Segmentation

Energy-based segmentation method consist of finding the optimal or best segmentation point according to an energy function. Consider that the grayscale image I consists of n voxels of side length . The spatial position and

the given intensity of voxel I are denoted xi and Ii, respectively. The segmented image ,where 0 stands for the background and 1 for object. Li is the segmentation label (0 or 1) of the voxel i. A segmentation L costs the energy E(L) The best-segmented image Lmin is the one that minimizes the energy E. In general, the energy E mainly consists of two terms: a local data fidelity term (Ev) and a non-local spatial regularization term (Es). This idea can then be formalized with proximity functions P that will quantify the penalty of assigning a particular voxel to object or to the background. The proximity function ranges from 1(closest) to 0 (furthest). Using proximity functions, the data fidelity term Ev of the segmentation energy can be defined as

$$v \cdot \sum_i ((1 - L_i) \cdot P_0(I_i) + L_i \cdot P_1(I_i)), \quad [11]$$

Figure 2. Data fidelity term formula

where P0 is the proximity to the background (0), and P1 the proximity to object(1). Using these proximity functions actually amounts to transforming the initial gray-scale image into a new gray-scale image which is sharper because the noise has been reduced. The complete energy function used is composed of two components:

$$E(L) = E_v(L) + r \cdot S(L), \quad [11]$$

Figure 3: Complete Energy based segmentation formula

here the spatial regularization term is Es = r.S(L), with S(L) the surface area of the segmented object and r (r ¿= 0) which has the dimension of the length. Including this term in the energy, it leads to the penalizing large interface areas: a voxel with a mixed gray value will be segmented such that the interface ice/background area will be minimized. The parameter r assigns a weight relative to the surface area term such that the smoothness of the segmented ice can be controlled. In principle, other non-local terms (e.g. curvature, gradient or shape penalties) could also be considered significant in the segmentation. The flexible definition of the energy function is known to be an asset of energy-based segmentation. The regularization term which minimizes the air/ice interface is of particular interest for materials like snow where sintering naturally tends to reduce the surface and grain boundary energy.

B. K- means Clustering

K-means clustering one of the popular unsupervised machine learning algorithm. The algorithm inputs include number of clusters and data set. The data set is a collection for each data point in the model. The algorithm initially chooses K centroids; This is randomly generated or randomly selected from the data set. To get the number of clusters, the Kmeans clustering algorithm is done for a range of K values and then the results are compared. K Means is a clustering algorithm. Clustering algorithms comes under unsupervised algorithm which means that is has no labeled data. It divides the data into clusters based on how the similar data is, where K refers to the number clusters in the model. Clustering algorithm easily adapts to new examples and guarantees convergence. The algorithm takes the unlabeled dataset as input, divides the dataset into k-number of clusters, and then it repeats the process until it finds the best clusters.

The value of k is pre-decided in this algorithm. The k-means clustering algorithm mainly performs two tasks:

1) It finds best K clusters or K centre points in an iterative process.

2) Designating each data point to the closest K Centre point. So we have each data point allocated to a particular cluster.

C. Convolutional Neural Network

Convolutional neural network needs a lot of data and computes resources to work efficiently for large images. CNN's are used for image classification as well as recognition because of their high accuracy Convolutional neural networks which are made up of neurons with learnable weights and biases. Each neuron receives many inputs and then takes a weighted sum over them, after which activation function is used. The very first layer detects simple and basic features like horizontal, vertical, and diagonal edges. The output of the layer is fed as input of the next layer in the model, which in turn extracts more complex features for example as corners and combinations of edges. Then we perform convolution by multiplying each pixel with weights. A CNN is usually

composed of many interconnected convolution layers. The final layer of a CNN is basically a classification layer, which takes the output of the final convolution layer as input (the higher convolution layers detect complex objects). In the beginning, the CNN starts off with random weights that can be assigned in a number of ways. During training, the neural network is provided with a large dataset of images annotated with their corresponding classes. It processes each image with its random values and then it compares the output with the input or image's correct label. If the output does not match the label, at the beginning of the training process, small adjustments are done to the weights of the corresponding neurons so that error reduces, and next time it sees the same image, its output will be a bit closer to the correct answer. These corrections or errors are done by a technique called backpropagation. Backpropagation is a way to find what correction should be done to minimize error. Every run of the entire training dataset is called an "epoch." The model goes through several epochs during training, which results in adjusting the weights in small amounts. After each epoch, the neural network becomes a bit better(which results in reducing the error)at classifying the training images.

As the CNN improves, the adjustments done to the weights become smaller and smaller. At some point, the network "converges," which means it has reached its peak. Pooling layer - The pooling layer is used to reduce the spatial volume of the input images after convolution. It is used between two convolution layers. If we only apply Convolutional layer without applying pooling or min/max pooling, then it will be computationally expensive. So, the max-pooling layer is very important and it helps in reducing the spatial volume of input image. Fully connected (FC) layer: Fully connected layer involves weights, biases, and neurons. It connects neurons in one layer to the neurons in next or another layer. It is used to classify images among different categories by training. Softmax/logistic layer: Softmax or Logistic layer is basically the last layer of CNN. It resides at the end of the FC layer. Logistic layer is used for binary classification which is mainly two outcomes (0 or 1) and softmax is for the multi-classification. Output layer: Output layer contains the label or the output which is in the form of one-hot encoded. After training the CNN, the developers use a test dataset to test the accuracy and see how it does on different dataset. The test dataset is a set of labelled images that are not a part of the training process. The output is compared to the actual label of the image. Essentially, the test dataset evaluates how good the neural network is in classifying images it has not seen before.

III. RESULTS AND DISCUSSION

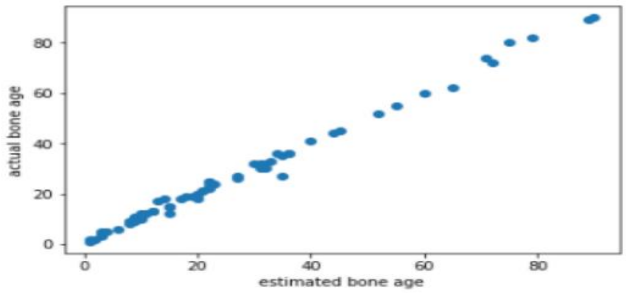

Figure 4: actual bone age vs estimated bone age

From the above graph, We can see that the dataset is spread through all ages. A lot of input images are in the range of 1 to 25.

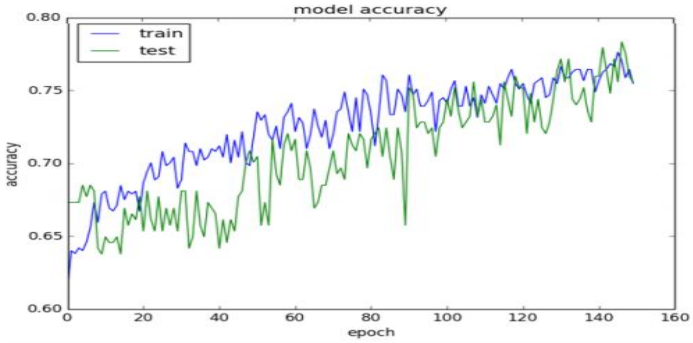

Figure 5: model accuracy.

Above image suggests that the accuracy keeps on increasing in epochs in both training and testing datasets, And after sometime it stays constant.

IV. CONCLUSION

After implementing energy based segmentation, K-means clustering is used to group similar type of data. Convolutional neural network is performed on the previous results which results to the predicted hand bone age. The results are more accurate as energy based segmentation was used to make the image more fine so that it can pick up more details. Convolutional neural network used gives better results as compared to Tanner white and GP method. The Convolutional neural network helps to find patterns and Regions of interests (Like horizontal edges, vertical edges and more). More experiments and comparisons are expected to fully study these issues. In the future, we will conduct a comprehensive and systematic study on this topic. We believe the performance of this method could be further improved. We can implement cryptography methods to improve the security of the assessment X-ray images and reduce the complexity factor. The proposed work can be extended using sand development of multi-factorial methods, involving hardware devices for easier bone age estimation ultrasound, MRI, IR images and low-dose CT technology,. We might also be able to increase the accuracy by separating the dataset into male and female cohorts. Girls generally begin and finish their puberty earlier than boys, and since we do not take this into account during training and testing we most likely lose some potential accuracy. In ROI classification, we select a uniform ROI size. However, different bone site should be investigated by different sizes. We can investigate this in future. This framework owns very high flexibility that a number of issues could have impacts on the final performance, such as the image pre-processing techniques, the parameters for multiscale and multi-direction decomposition, the specific network architecture, etc.

ACKNOWLEDGEMENTS

It gives us immense pleasure to express our deep gratitude and sincere thanks to our major project guide Prof. Kumud Wasnik for her valuable and useful support and comments towards this project. Her continuous encouragement promoted us to do our task rigorously.

V. REFERENCES

[1] Erik Westerberg, “AI-based Age Estimation using X-ray Hand Images”, June 8, 2020.

[2] Ms. K. Meenakshi Prabhakaran, Ms. Hemalatha, Ms. Mahima R.V., Ms. Maitri Manjunath Bhat, Handbone Age Estimation Using Energy Based Segmentation, Sep 03, 2020. [3] R. K. Bull, P. D. Edwards, P. M. Kemp, S. Fry, and I. A. Hughes, “Bone age assessment: A large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods,” 2017.

[3] P. Thangam, K. Thanushkodi ,T. V. Mahendiran: “Comparative Study of Skeletal Bone Age Assessment Approaches using Partitioning Technique” International Journal of Computer Applications Volume 45– No.18, May 2012.

[4] Xiaohong Liu, “Bone Age Assessment by Deep Convolutional Neural Networks Combined with Clinical TW3-RUS”.

[5] Xun chen, Chao Zhang, Yu Liu,”Bone Age assessment with X-ray images based on contourlet motivated Deep Convolutional Networks”, IEEE, August 2018.

[6] Ravi Kaushik, Shailender Kumar, “Image Segmentation using Convolutional neural networks”, IJSTR, 11 November, 2019

[7] Shweta M Madiwal, Mahesh Prasanna K ,”Hand bone age estimation using energy based segmentation”, 6, June 2017

[8] D. Giordano, R. Leonardi , F. Maiorana , G. Scarciofalo , and C. Spampinato,”Epiphysis and Metaphysis Extraction and Classification by Adaptive Thresholding and DoG Filtering for Automated Skeletal Bone Age Analysis”,Conference of the IEEE EMBS, France/ 2007

[9] Rijad Saric, Jasmin Kevric, “Evaluation of Skeletal Gender and Maturity for Hand Radiographs using Deep Convolutional Neural Networks”,2019.

[10] Pascal Hagenmuller, Guillaume Chambon,Bernard Lesaffre, Fred´ eric ´ Flin, “Energy-based binary segmentation of snow microtomographic images”, May 2013.