RESEARCHARTICLE OPENACCESS

Using Machine Learning Techniques For Bone Age Assessment

Nur Zakiroğlu* Ali Güneş** Metin Zontul** Ahmet Gürhanlı**
**(Department of Computer Engineering, Istanbul Aydın University, Istanbul,Turkey*
***(Department of Computer Engineering, Istanbul Aydın University, Istanbul, Turkey*
Corresponding Author : Nur Zakiroğlu

ABSTRACT
Bone age is an effective indicator for the diagnosis of various diseases. A person has two ages, chronological age and bone age. Chronological age is the actual age determined from the date of birth of the person. Bone age describes the degree of maturation of the person's bones. Many diseases, such as growth disorders, chromosomal disorders, endocrine disorders, and endocrinological problems, can be discovered by finding inconsistency between bone age and chronological age. In this article bone age is determined by applying image processing techniques on people hands x-rays with techniques of machine learning. The hand x-rays image data of 2500 children were studied. CNN VGG16 and CNN Inception V3 regression algorithms were applied to these data and then the results were evaluated.
Keywords: Bone Age, Machine Learning, Image Processing , Artifical Intellegence

---

Date Of Submission: 28-07-2019 Date Of Acceptance: 10-08-2019

I. INTRODUCTION

Bone age assessment is often performed not only in pediatric endocrinology, but also in pediatric orthopedics. Bone age is an effective indicator for the diagnosis of various diseases and the timing of treatment. The aim of bone age assessment is to assess growth and maturity and to diagnose and manage pediatric diseases. Therefore, the accuracy of bone age assessment is very important. Recently, various computerized systems for bone age assessment have been developed. This study describes the progress in evaluation methods and clinical applications of the bone age. In this article, a person's bone age is determined by using machine learning and image processing techniques on person’s x-ray image of his left hand. A child has two types of age, chronological age and bone age.

The chronological age is the actual age determined from the date of birth of the child. Bone age defines the degree of maturation of the child's bones. The changes in the development of the human skeleton are basically similar, because the development process of each bone goes through the same stages.

At each stage, the bones have their own characteristics. Therefore, bone age is a more accurate way of reflecting individual growth development level and maturation compared to chronological age [1].

II. IMAGE PROCESSING ON X-RAY IMAGES

By applying image processing methods on images, images quality are improves and we can obtain information from images. A digital x-ray is used to capture images, scan an image, or store the image in digital format such as jpeg, png, so it is useful in diagnosing patients' illness. However, digital x-ray images include gaussian noise or salt and pepper noise, so sometimes they do not give an X-ray image clearly. Therefore, image processing methods are important to remove n oise in the image to improve the quality of the images [2].

Figure 1: An example of the application of noise reduction and image smoothing in an x-ray hand image.

2.1. Noise Reduction

Noise can be defined as unwanted pixels that affect the quality of the image. Noise can be written as:
$f(x, y) = g(x, y) + \eta(x, y)$

f (x, y) is the original image, g (x, y) is the output image and η (x, y) is the noise model. There are different types of noise. Salt and pepper noise is one of the most common types of noise that can be found in x-ray images [3].

In this study, a median filter was used to reduce salt and pepper noise while preserving image edges and sharpness. The median filter is also used to reduce noise from the image while maintaining the edges and sharpness of the image. The median filter takes each pixel in the image and controls how different it is from the neighboring pixels, if "very different", its value is replaced with the middle value of the surrounding pixels [3]. Figure 1 shows an example of the application of noise reduction and image smoothing in an x-ray hand image.

2.2. Image Edge Detection

In edge detection, pixels are reduced and images are preserved. Edge detection is a method of identifying points at which the image brightness changes sharply or becomes more blurred. Edge detection can be defined as the discovery of lines marking the boundary of the image and separated from other places or objects in the digital image.

2.3. Feature Extraction and Selection

After correcting the image and fixing the edges of the hand bone, the proposed system is to extract useful and distinctive features of the hand bone image. Feature extraction is the main step in various image processing applications. A combination of different sets of properties is used, such as the properties of the wavelet transform, the properties of the Curvelet transform and other textural features [4].

III. MACHINE LEARNING

Machine learning is a data analysis technique that teaches computers to do what people and animals do naturally in daily life. It is an application of artificial intelligence (AI), which makes software applications more accurate in predicting results without being explicitly programmed [5].

3.1. Learning With CNN

A neural network is a system of artificial "neurons" connected to each other, exchanging data between them. Connections have numerical weights set during the training process, so they will work correctly when presented with an image or pattern to recognize a properly trained network. The network consists of a layer of on neurons that perceive multiple features.

Each layer has many neurons that respond to different input combinations than the previous layers. A CNN consists of one or more layers having a sub-sampling layer, usually followed by one or more fully connected layers as in a standard neural network. In the conventional pattern / image recognition model, a manually designed feature extractor collects relevant information from the input and eliminates trivial variability [6].

3.2.Used Data Set

The data set for bone age determination was developed with the contributions of Stanford University, Colorado University and the University of California. The data sets were published by the universities on their web pages and received from them. The data set used consists of children's hand x-rays. Hand x-rays of boys and girls aged 0-18 years were used for bone age determination. Image data sets of 2500 x-rays were fully trained and showed good results. Convolutional Neural Networks (CNNs) were used for image classification tasks. A lot of data and time is needed to train CNNs. However, sometimes the dataset may be limited and may not be sufficient to train a CNN from scratch. In such a scenario, it helps to use a previously trained CNN on a large dataset.

3.3. Transfer Learning

In transfer learning, a basic network is first trained on a basic dataset and task, and then retargeted learned features or transferred to a second target network to be trained for a target dataset and task [7]. The VGG-16, a 1-layer network trained in Imagenet, used as pre-trained CNN. Then, Inception V3, a 1-layer network trained on Imagenet on the same datasets, was used also as a pre-trained CNN and compared with VGG16 results.

IV. MODELLING

Deep CNNs consist of varying convolution and pooling layers to learn the layered hierarchy from input images, followed by fully linked classification layers, which can then be trained with property vectors from previous layers. Many innovative deep neural networks and new training methods have shown impressive performance for image classification tasks. Inspired by a common family of VGG models, it is implemented as a deep convolutional neural network with regression output.

The VGG module consists of two convolution layers with Exponential Linear Unit (ELU) activation function, cluster normalization and maximum pooling. The input image is passed through a network of three VGG blocks followed by three Fully Connected layers [8].

VGG blocks consist of 64, 128, 256 convolution layers, respectively. The model is trained with the Mean Square Error Loss function (MSE) using the Adam optimizer:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \widehat{y_i}|$$

4.1. VGG16 - Convolution Network for Classification and Detection

VGG16 is a convolutional neural network model introduced by Simonyan K. and Zisserman A. of in their paper "Very Deep Convolution Networks for Large Scale Image Recognition " by Oxford University.

The model achieved the first 5 test accuracy of 92.7% in ImageNet, the data set of more than 14 million images of 1000 classes[8]. Figure 2 shows the VGG16 architecture by replacing large core-size filters (in the first and second evolutionary layers respectively) with several 3 x 3 core-size filters one after the other.

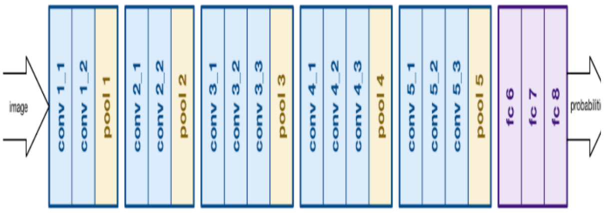

Figure 2: The VGG16 architecture

4.2. Inception V3

Inception v3 is a widely used image recognition model that has been shown to achieve greater than 78.1% accuracy in the ImageNet dataset. The model itself consists of symmetrical and asymmetrical building blocks, including convolutions, average pooling, maximum pooling and fully interconnected layers. Inception-v3 is a convoluted neural network that has been trained on more than a million images from the ImageNet database.

The network is 48 layers deep and can split images into 1000 object categories such as keyboards, mice, pencils and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 299 to 299 [9].

V. RESULTS

The same image preprocessing and feature selection methods were applied to CNN transfer learning models before feeding them into datasets, using pre-trained weights with VGG16 and Inception V3. The mean absolute error of the months (MAE) was used to measure the difference between the estimated age and the actual age.

Twenty percent of the data set was chosen as a random validation set and obtained using two models on MAE validation sets. In VGG16 Transfer Learning MAE (Absolute Error) result is 14.43, while Inception V3 Transfer Learning MAE result is calculated as 37.29. It is clear that the Incepition V3 transfer model performs worse than the VGG16 transfer model. Both models were trained using Adam opitmazation.

In the training process, pre-trained weights using CNN transfer learning were selected as starting weights, the model was re-trained on new data sets and the best model weight was determined. In transfer learning, both pre-trained VGG16 and pre-trained Inception V3 and both weights from ImageNet were selected as initial weights, and the learning status was compared between the two methods.

As a result, the model trained with the CNN transfer learning method, VGG16 was chosen as the best method with pre-trained weights up to 86% accuracy when used in 2500 Hand X-Ray training.

Figure 3 shows the difference between the predicted age and the actual age on the data from the training result.

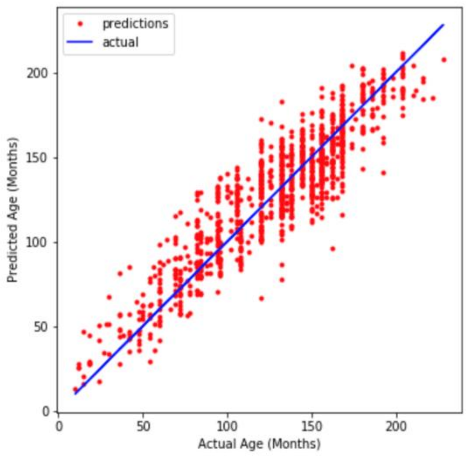

Figure 3: The difference between the predicted age and the actual age in VGG16

When the pre-trained weights with Inception V3 are used in 2500 Hand X-Ray training, the accuracy reaches up to 63%. As a result of this method, accuracy has been accepted as low. Figure 4 shows the difference between the predicted age and the actual age on the result of training.

ISSN : 2248-9622 Vol. 9,Issue 8 (Series -II) Aug 2019, pp 16-20

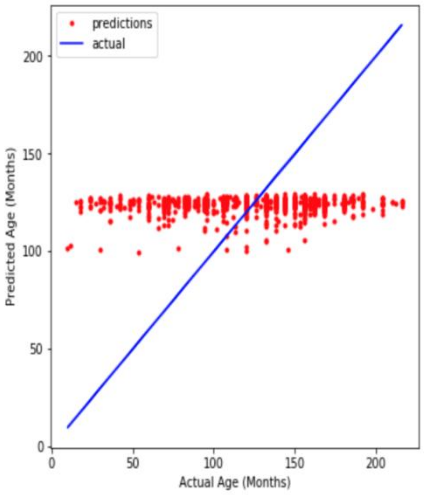

Figure 4: The difference between the predicted age and the actual age in Inception V3

Guided by CNN Tranfser learning and getting the best results, the VGG16 model showed the best results in estimating bone age using image processing techniques on image data sets. A few of the test data are selected below. As shown in Figure 5, there is little difference between Actual Age and predicted Age.

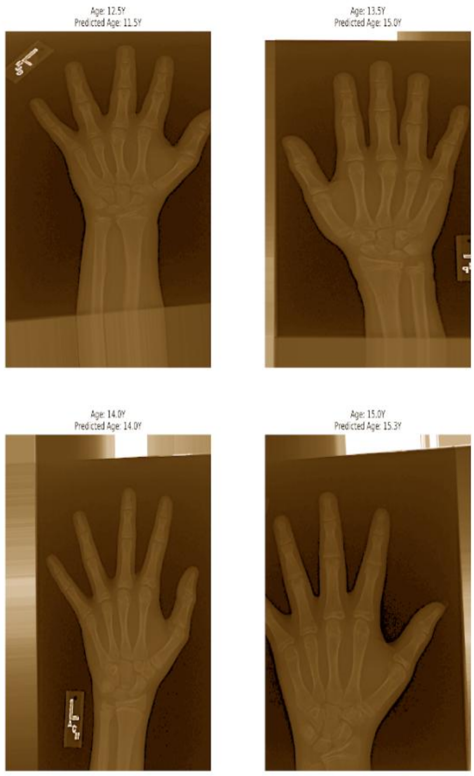

Figure 5: Examples of VGG16 results that shows difference between actual and predicted age

The Inception V3 model, which is best cured by CNN Tranfser learning, showed lower results compared to VGG16, which predicts bone age using image processing techniques on image data sets. A few of the test data are selected below. As seen in Figure 18, the difference between Actual Age and Estimated Age is slightly higher.

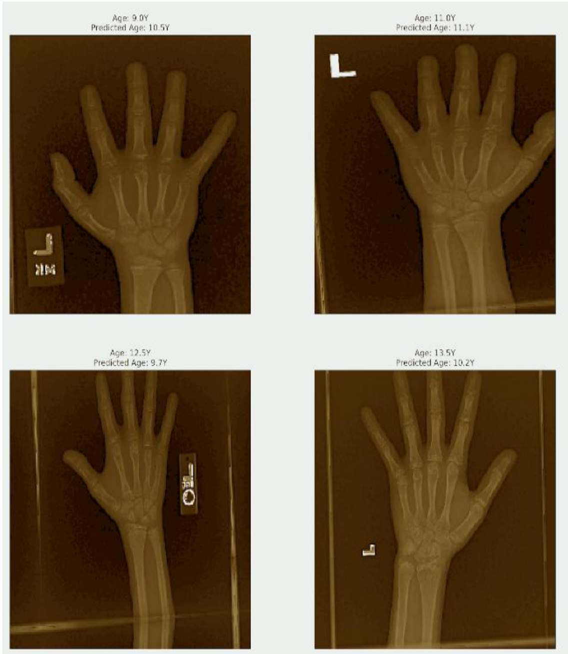

Figure 6: Examples of Inception V3 results that shows difference between actual and predicted age

VI. .DISCUSSION

In this study we used two type of Transfer learning techniques to predict the age of bone by entered x-ray hand image. Images processing techniques was applied on 2500 x-ray hand image and best result was from VGG16 CNN Transfer learning. For better result we can try the training on more data set with applying the same techniques on more images. and another way is to compare between typical CNN learning result on bone age detection and CNN Transfer learning results that we get. VGG16 training take a more time than Inception V3 , with much better results. We can try to apply another transfer learning techniques and compare them with our results.

REFERENCES

[1]. De Sanctis V, Di Maio S, Soliman A.T' Rai, G, Elalaily R, Millimaggi G.(2019).Hand X-ray in pediatric endocrinology: Skeletal age assessment and beyond. DOI:10.4103/2230-8210.145076

[2]. Chikhalekar A.T.(2016).Analysis of Image Processing for Digital X-Ray. International Research Journal of Engineering and Technology (IRJET)

[3]. Jassim, F. A. (2003). Kriging interpolation filter to reduce high density salt and pepper noise. World of Computer Science and Information Technology.

[4]. Al-Ayyoub M, Hmeidi I, Rababah H. (2013).Detecting Hand Bone Fractures in X-Ray Images. Jordan University of Science and Technology
[5]. Iliadis N, (2017).Machine learning: the power and promise of computers that learn by example. ISBN: 978-1-78252-259-1,The Royal Society.
[6]. O'Shea K.T, Nash R.(2015).An Introduction to Convolutional Neural Networks Article. arXiv:1511.08458
[7]. Torrey L, Shavlik J.(2009),Transfer Learning. University of Wisconsin, Madison WI, USA,DOI: 10.4018/978-1-60566-766-9.ch011
[8]. Liu I, Li X, Luo P, Loy C.C, Tang X,(2015),Semantic Image Segmentation via Deep Parsing Network.
[9]. Barratt S, Sharma R (2018),A Note on the Inception Score,arXiv:1801.01973

Nur Zakiroğlu" Using Machine Learning Techniques For Bone Age Assessment"International Journal of Engineering Research and Applications (IJERA), Vol. 09, No.08, 2019, pp. 16-20