ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AUTOMATIC BONE AGE ASSESSMENT USING HAND X-RAY IMAGES

NOOR MUALLA[1], ESSAM H. HOUSSEIN[2], AND M. R. HASSAN[1]

[1]Computer Science branch, Mathematics Department, Faculty of Science, Aswan University, Egypt

[2]Department of Computer Science, Faculty of Computers and Information, Minia University, Egypt

Corresponding author: Essam H. Houssein (e-mail: essam.halimr@mu.edu.eg).

ABSTRACT

The accurate age estimation of human is crucial in several cases and fields. Bone Age Assessment (BAA) can be an effective method for human age estimation for live and dead human. Several manual methods have been proposed to achieve this task; however, they are time consuming and error-prone. In this paper, an automated bone age assessment system is proposed based on the concept of transfer learning for feature extraction. In the proposed system, the hand X-ray images are preprocessed, and the discriminant features are extracted using a pre-trained, fine-tuned deep neural networks (AlexNet and ResNet-101). Finally, the age group of the hand X-ray image is determined using a number of classification models including decision tree, k-nearest neighbour, linear discriminant, and support vector machine. The proposed system is assessed using the RSNA Bone Age dataset. The obtained results have shown that the ResNet-101-based features are more effective than the corresponding AlexNet-based features. In addition, the decision tree classifier is better than the remaining classifiers with classification accuracy up to 100%.

Keywords: *Classification, Bones Age Assessment, Deep Learning, Transfer Learning, X-ray Images.*

1. INTRODUCTION

Machine learning approaches have been widely employed in the medical field over the past years in order to perform several tasks. One of the important tasks that have attracted an increasing attention of many researchers is the accurate age estimation of dead or live human based on medical images. The human age estimation can be classified into two categories: Age estimation for live human and Age estimation for dead human. In the first type, the human age can be estimated using different types of radiograph images such as X-ray, MRI, CT and DICOM. On the second type, the human age can be estimated through manually analyzing the characteristics of the human body remains [1].

Age estimation for live human is critical in many fields and due several reasons. For example, the accurate age estimation can help in criminal identification and law enforcement. In addition, accurate age estimation is necessary for military service or sportive championships that need a certain range of age, particularly when we know that many parents attempt to register their children with wrong birthdate [2]. In addition, accurate age estimation can play a great role in the medical diagnosis process.

Several manual bone age estimation approaches have been proposed, particularly for dead people. In these approaches, the age as well as other information can be obtained from the dead persons' skeleton bones such as gender, height, cause of death, etc. However, in some cases such as earth calamities, bomb blasts, and tsunamis, the forensic researchers may get skeleton remains or half decomposed body. Therefore, many research efforts have been performed for age estimation based on parts of the bones such as teeth [3-6], skull [7, 8], foot bones [9, 10], knees [11, 12], etc. However, manual bone age estimation method suffers from a set of disadvantages such as subjectivity due to the dependence on human factor, error-prone, time consumption, the need for special instruments, etc [1].

Human hand bones have many significant properties that can be employed effectively in human age estimation. Many methods have been proposed for human age estimation based on hand image. However, these methods involve intensive image processing in order to prepare the images for the consequent stages. In addition, in order to compute the hand lengths, these methods usually perform a segmentation process which is a very challenging process. Moreover, the success of the future extraction step highly depends on the success of the segmentation step. The deep learning can be used effectively to perform the feature extraction step without the need for the segmentation process by taking the hand image as input and returning the hand features that can be used in the age estimation process as output.

In this study, an automatic bone age assessment system is presented that depends on X-ray images. The objective of the proposed system is to determine the age group of a person using its hand X-ray images among different seven age groups. The proposed system includes three main steps: Image preprocessing, feature extraction and classification. The feature extraction step is performed using two convolutional deep neural networks namely AlexNet and ResNet-101 while the classification step is performed using a number of well-known classifiers including decision tree (DT), k-nearest neighbor (k-NN), linear discriminant (LD), and support vector machine (SVM).

The remaining sections of this paper are organized as follow: Section 2 reviews and summarizes different bone age estimation methods. Section 3 describes the proposed bone age assessment method in detail. Section 4 contains the implementation details and results analysis. Finally, the study is concluded, and the future work is recommended in Section 5.

2. RELATED WORKS

Early research efforts have been performed in order to achieve accurate bone age assessment. Several manual methods that depended on left hand bones have been proposed including Greulich and Pyle Method (GP) [13] and Tanner and Whitehouse Method (TW) [14, 15]. However, these methods were time consuming and the output results varied due to the user's subjectivity. Therefore, many researchers attempted to develop fully automated systems for bone age assessment. One of the first attempts called Fels Longitudinal Study (FELS) was proposed by A.F. Roche et al. [16]. In the FELS method, the age is estimated based on the scores/grades of each bone in the left-hand image. In addition, it can predict the error margin in the estimated ages. However, the authors assumed that the input images must be clear and distortion-free images to obtain good age estimation accuracy.

Another semi-automatic bone age estimation method that depends on X-ray images of hand bones has been proposed by David J. Michael and Alan C. Nelson [17]. The proposed system called HANDX Software System. This system involves three main steps: preprocessing, image segmentation, and measurement. In this system, the background pixels, soft tissues pixels and bone pixels are separated through constructing and segmenting the histogram of hand images. Then, the measurements are computed for each bone and manually compared to the TW2 and GP atlas.

Another early automated bone age assessment method that depends on phalangeal bones has been proposed by Eva Pietka et al [18]. In the proposed method, the background pixels are moved, the boundaries of the bones are identified, and the rotation process is applied on demand. Then, Phalangeal ROIs are determined using the sobel gradient. Finally, a set of measurements are computed and compared to length table of phalangeal [19] for Bone Age Assessment. The proposed system achieved accurate age assessment in 94% of the cases.

Tanner and Gibbons [20] have proposed a computerized bone age estimation system called CASAS. In the proposed system, the X-ray images are digitized, and each bone is determined. Then, the age assessment is performed automatically based on the Fast Fourier Transform (FFT). Some researchers have conducted a comparison between the CASAS system and the manual TW method, and the obtained results have shown the superiority of the CASAS system.

Gross et al. [21] have introduced an automated bone age assessment system based on hand wrist radiograph. The proposed system starts with extracting ten measurements which further reduced to seven using linear regression analysis. The age estimation process is performed using a decision system that depends on a neural network. However, the proposed method has not employed the morphological features used in the GP and TW methods. Also, the performance of the proposed system is nearly the same as the GP manual method.

Sato et al. [22] have presented an automated bone age assessment system called CASMAS for Japanese children based on the third digit. The proposed system employs the proximal, middle, and distal epiphyses of the third finger in the age assessment process. The conducted experiments have shown that the proposed system gives acceptable results for the children of age range between 2 and 15 while the accuracy of the system decreases for those above 15.

Hsieh et al. [23] have suggested another bone age assessment based on the third digit of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

left-hand X-ray image. The radiograph is rotated using thresholding methods and other heuristic searches. The phalangeal region of interest (PROI) is segmented using Gabor filter in addition to canny edge detector. The extracted features include geometric indicators as well as the information of the epiphysis shape of the distal phalanx. Three neural networks are employed to perform the bone age assessment namely, back-propagation, radial basis function, and support vector machine. The experimental results have shown that the support vector machine is the best with 80% accuracy.

Thodberg el al. [24] have presented a bone age assessment system called BoneXpert. It depends on shape-driven active appearance in addition to the TW RUS-based method (radius, ulna and short bones). The bones contours are rotated and scaled using the Gabor filters. The linear regression analysis is used to select the most discriminant 30 coefficients which are used as the input features in the active appearance model. The conducted experiments have shown that the proposed system achieved reasonable results with accuracy about 0.42 using the GP method and 0.8 years using the TW2 method. An extension for this work is provided in [25] where the old work covered the ages up to 17 years for boys and 15 years for girls while the new work covers the ages up to 19 years for boys and 18 years for girls.

Lee et al. [26] have proposed a fully automated deep learning pipeline in order to perform the different tasks toward and including the bone age assessment. Among the performed tasks, the proposed system segments the ROIs and applies the necessary preprocessing for the used radiographs. The proposed system employs the pre-trained, fine-tuned convolutional neural networks that have been suggested for the ImageNet challenge. The experimental results have shown that the proposed system has 57.32 and 61.40% accuracies for females and males, respectively.

Spampinato et al. [27] have introduced another bone age assessment system based on the deep learning techniques. The obtained results have shown that discrepancy between the performance of the proposed system and the state-of-are manual methods is about 0.8 years. In addition, the proposed work has conducted a comparison between the traditional hand-crafted features and the deep learning-based features in the medical field which has shown the effectiveness of the later type.

Chu et al. [28] have proposed a bone age assessment system that depends on deep convolution network. The proposed system consists of two stages. In the first one, the mask of bones is generated using a U-Net convolution network with pre-trained VGG16 as the encoder. In the second stage, the original images are fused with the generated masks in order to obtain the ROIs from the hand bone images which are used as input to multiple-output convolutional neural network to perform the bone age assessment. Finally, the age regression problem is transformed into K-1 binary classification problem. The proposed system is evaluated using the RSNA2017 Pediatric Bone Age dataset and the obtained results have shown that the proposed system has a mean absolute error (MAE) of 5.98 months.

Kaur and Mann [29] have proposed a segmentation method for bone parts using an enhanced point distribution algorithm based on contours. The segmented bone parts can be used effectively in the bone age assessment process. The implementation of the proposed segmentation algorithm has shown the ability of the proposed algorithm to segment bone parts from the left hand X-ray image with high accuracy.

Mutasa et al. [30] have introduced a bone age assessment based on a customized, purpose-built deep convolutional neural network not pre-trained, fine-tuned deep neural networks. In the proposed system, the linear regression output is used. The conducted experiments over a dataset that consists of 10,289 images have shown that the proposed system has a promising performance in terms of mean absolute error.

Tajmir et al. [31] have suggested another bone age assessment system that depends on deep convolutional neural networks. The proposed system is evaluated using 280 age and gender-matched bone radiographs with ages range between 5 and 18. The obtained results have shown that the proposed system has 68.2% overall accuracy and 98.6% accuracy within 1 year.

3. THE PROPOSED WORK

In this section, a bone age assessment system is proposed based on hand X-ray images. The proposed system aims to determine the age group of persons using their hand X-ray images. As shown in Figure 1, the proposed system consists of three main steps: image preprocessing, feature extraction, and classification. In the image preprocessing step, the hand X-ray images are converted into the RGB color model and resized to a certain size. Additionally, a data augmentation process is applied to enlarge the size of dataset. In the feature extraction step, the concept of transfer

Journal of Theoretical and Applied Information Technology
31st January 2020. Vol.98. No 02

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

learning is adopted to perform this task. Two pre-trained deep neural networks, namely AlexNet and ResNet, are employed to extract the discriminant features. Finally, a number of classification models are used to perform the classification task including Decision Tree (DT), K-Nearest Neighbor (K-NN), Linear Discriminant (LD), and Support Vector Machine (SVM). The detailed descriptions of the different steps are given in the following subsection.

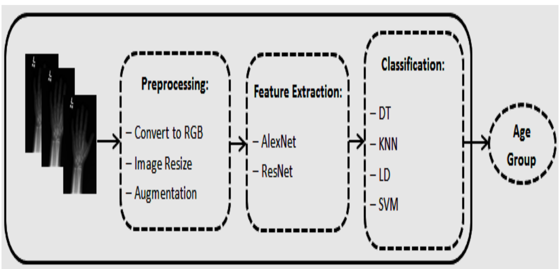

Figure 1: The block diagram of the proposed bone age assessment system

3.1. Image preprocessing

In this step, three main operations are performed to make the hand X-ray images ready for subsequent steps. First, hand X-ray images are converted into RGB color model. Then, hand X-ray images are resized to fit the requirements of the used deep neural networks. The size of hand X-ray images is set to 277*277 with AlexNet while the images' size is set 224*224 with ResNet. Finally, a data augmentation step is performed to build a large dataset that is suitable for deep neural networks. In this step, three operations are performed namely, translation and reflection. In the translation operation, the hand X-ray images are randomly shifted along the X-axis and Y-axis with a shift value bounded by the interval [-30, 30]. In the reflection operation, the hand X-ray images are mirrored along the vertical access. In the rotation process, the hand X-ray images are rotated with a random angle bounded by the interval [5°, 35°]. An example for the data augmentation process is shown in Figure 2. The left column contains the original hand X-ray images. The second column contains the translated versions. The third column contains the mirrored versions. The last column contains the rotated versions.

3.2. Feature extraction

Convolutional Neural Networks (CNN) is a popular architecture for deep neural networks that achieved many breakthroughs in many fields including machine learning and computer vision. CNNs can successfully accomplish their work without being affected by tilting, translation, and scaling [32]. CNNs usually include three layer types: convolutional layer, pooling layer, and fully connected layer (See Figure 3.).

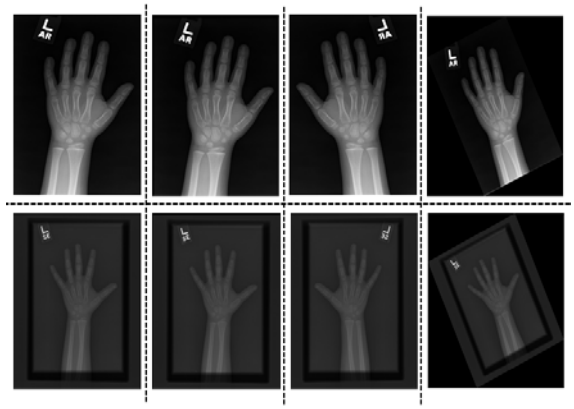

Figure 2: Data augmentation process for hand X-ray images

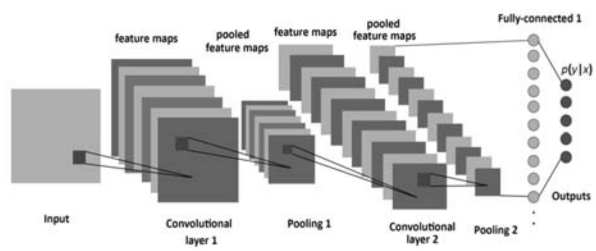

Figure 3: The general architecture of convolutional neural networks

The role of convolutional layer is to compute the weighted sum, to add the bias value to the weighted sum, and to apply an activation function called the rectifier linear unit (ReLu), which is defined using Eq. 1, on the addition result. On the other side, the objective of pooling layers is to manage the over-fitting by decreasing the number of features obtained from the convolutional layer. Finally, the fully connected layers aim to gathering all the feature of descriptor to be classified using the last layer [33].

$$ReLu\ (x) = \max(0, x) \tag{1}$$

In the proposed work, two well-known pre-trained convolutional neural networks called AlexNet [34] and ResNet-101 [35] are used to perform the feature extraction step using the concept of transfer learning. Generally, transfer learning is used for a number of reasons. First, training a CNN from scratch using random initial values is difficult due to the absence of large datasets. Hence, using the weights of a pre-trained net as initial values can be useful in addressing

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

many of the problems in hand. Second, training a very deep network from scratch is a time-consuming process that needs sophisticated machines with expensive GPUs. Finally, there is no clear theoretical guidance that can help in selecting the appropriate topology, training method, parameter values, etc [36].

3.2.1 Feature extraction using AlexNet

AlexNet is a popular CNN that was proposed by Krizhevsky et al. [34] to compete in the ILSVRC-2010 challenge for classifying the ImageNet database. It contains five convolutional layers, three fully-connected layers, as well as max-pooling layers. All of the eight layers need to be trained. In AlexNet, the overfitting problem is addressed using a number of ways including normalizing the local response, data augmentation, and the dropout approach in which the output of hidden neurons is set to zero with a probability 0.5. The dropout process is performed on the first two full-connected layers. In the feature extraction step, the last three layers of the original AlexNet are freezed and replaced with other three layers that suit the classification problem in hand. The eliminated layers are the last fully-connected layer, the softmax layer, and the output layer. The features are obtained from the last fully-connected layers after completing the training process using our dataset. The length of each feature vector is 4096.

3.2.2 Feature extraction using ResNet-101

Deep residual networks are extremely deep architectures that have achieved high accuracy and good convergence behavior in many recognition and classification problems [37]. Deep residual networks have been proposed to address the problem of accuracy degradation that occurs when deep networks start to converge. The idea behind the residual learning is shown in Figure 4. Simply, each few stacked layers fit a residual mapping rather than directly fitting the required underlying mapping. In other words, if the required underlying mapping is denoted by *H(x)*, the stacked nonlinear layers fit another mapping which is defined as $F(x) = H(x) - x$. Hence, the original mapping is reformulated as $f(x) + x$ which can be achieved using feed forward neural networks and shortcut connections [37].

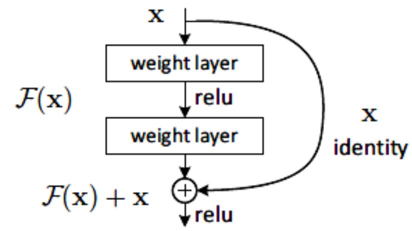

Figure 4: The idea of residual learning

ResNet is a very deep network which is designed and built based on the principle of residual learning [37]. It is the winner of ILSVRC 2015 in image classification challenge. Several variants have been designed that belongs to the ResNet family including ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. The number in the network name denoted the number of layers included in the network. In the proposed work, the pre-trained residual network ResNet-101 is used to perform the feature extraction step. The length of the resulting feature vector is 2048.

3.3. Classification

In the proposed bone age assessment system, the classification step is done using a number of well-known classification models. The goal of the used classification models is to determine the age group of a person based on the feature vector extracted from his hand X-ray image. The used classifiers include decision tree (DT) [38], K-nearest neighbor (K-NN) [39] using Euclidian distance and K is set to 1, linear discriminant (LD) [40], and support vector machine (SVM) [41] using different kernel functions.

4. IMPLEMENTATION AND EXPERIMENTS

This section includes the implementation details of the proposed bone age assessment system. Additionally, it describes the used dataset. Moreover, a number of experiments are conducted to evaluate the proposed system in terms of a number of performance metrics.

4.1. Dataset description

The used dataset is the RSNA Bone Age dataset [42]. The proposed system is evaluated using 1684 hand X-ray images that belong to persons of different genders and different age groups. A sample of the hand X-ray images form the used dataset is shown in Figure 5.

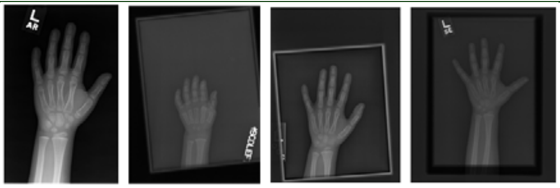

Figure 5: A sample from the used dataset

In the proposed bone age assessment system, eight age groups are considered. The details of the different age groups and the number of images contained in each group are shown in Table 1.

Table 1: The details of the used dataset

Class	Age Rang	Num. of Images
A	[0-9]	13
B	[10-19]	63
C	[20-29]	150
D	[30-39]	166
E	[40-49]	272
F	[50-59]	415
G	[60-69]	253
H	>=70	352

4.2. Implementation and experimental results

The proposed automated dental age estimation is implemented using Matlab 2018a. The performance of the proposed system is assessed using the following performance measures:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F - Measure = \frac{2TP}{2TP + FP + FN} \quad (6)$$

Where TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative. The used classifiers are evaluated using the 10-fold cross validation approach. The DT classifier is implemented using max-split is equal to 20. The K-NN classifier is implemented using K=1. The standard SVM works on two classes only while the proposed work needs multi-class classifier. So, a modified version of the standard SVM called Multi-Class Support Vector Machine [43] is employed for building a classifier that is capable of differentiating among several classes. Moreover, several kernel functions are used with the SVM classifier including Gaussian kernel function, cubic kernel function, and quadratic kernel function. AlexNet is trained with the following settings: the patch size is equal to 5, the number of epochs is equal to 6, and the learning rate is equal to 1*e-4. Similarly, The ResNet is trained using the following settings: the patch size is equal to 10, the number of epochs is equal to 6, and the learning rate is equal to 1*e-4. All the experiments are executed using GPU NVIDIA GE FORCE 920M 4 GDDRAM. The obtained results are shown Tables 2-12. In addition, a summary for the obtained results is shown in Table 13.

Based on Table 13, it is observed that the different classification model has very acceptable performance regarding the used performance metrics. However, using AlexNet based features, it is noticed that the DT classifier is the best regarding the accuracy, specificity and precision performance metrics while the K-NN classifier is the best in terms of the recall and F-measure performance metrics. On the other side, the SVM classifier with quadratic kernel function is the worst regarding the different performance metrics. Using the ResNet based feature, it noticed that the DT classifier is the best in terms of the different performance metrics while the SVM classifier with quadratic kernel function is still the worst. By comparing the AlexNet based features to the ResNet based features, it is observed that the performances of the different classification models are better with ResNet based features except the LD classifier that has failed with the ResNet based features due to the limited computing resources.

Unlike previous related works, the proposed system performs minor image processing on the used images where the deep neural networks can work directly on the raw pixels of images and returns useful discriminating features that can be used in human age estimation. In addition, the results achieved by the proposed system are better than those of previous works where the classification accuracy in the proposed work has reached 100%.

Table 2: The performance of the proposed system using AlexNet based features and DT classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	100	98.529	99.518	99.605	99.432
Specificity	100	100	99.935	100	100	99.842	99.86	99.699
Precision	100	100	99.338	100	100	99.518	99.213	98.87
Recall	100	100	100	100	98.529	99.518	99.605	99.432
F-Measure	100	100	99.668	100	99.259	99.518	99.409	99.15

Table 3: The performance of the proposed system using AlexNet based features and K-NN classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	100	98.529	99.518	99.605	98.86
Specificity	100	100	99.935	100	99.929	99.685	99.93	99.849
Precision	100	100	99.338	100	99.628	99.041	99.605	99.429
Recall	100	100	100	100	100	99.518	99.605	98.864
F-Measure	100	100	99.668	100	99.814	99.279	99.605	99.146

Table 4: The performance of the proposed system using AlexNet based features and LD classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	100	98.529	99.036	99.605	99.432
Specificity	100	100	99.938	99.868	99.858	99.764	100	99.775
Precision	100	100	99.338	98.809	99.259	99.275	100	99.15
Recall	100	100	100	100	98.529	99.036	99.605	99.432
F-Measure	100	100	99.668	99.399	98.893	99.155	99.802	99.291

Table 5: The performance of the proposed system using AlexNet based features and SVM classifier and quadratic kernel function

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	89.744	95.767	90	87.149	90.931	94.779	89.328	89.867
Specificity	100	99.959	99.631	99.78	97.71	95.745	98.3	97.823
Precision	100	98.907	95.972	97.748	88.439	87.928	90.254	99.164
Recall	89.744	95.767	90	87.149	90.931	94.779	89.299	89.867
F-Measure	94.595	97.312	92.89	92.145	89.668	91.225	89.774	94.287

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6: The performance of the proposed system using AlexNet based features and SVM classifier and cubic kernel function

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	100	98.899	99.036	99.736	99.242
Specificity	100	99.938	100	100	99.811	99.842	99.93	99.725
Precision	100	98.438	100	100	99.018	99.516	99.603	98.961
Recall	100	100	100	100	98.897	99.036	99.735	99.242
F-Measure	100	99.213	100	100	98.957	99.275	99.669	99.101

Table 7: The performance of the proposed system using AlexNet based features and SVM classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	89.744	100	100	100	98.162	100	99.605	98.295
Specificity	100	100	100	100	100	98.949	100	100
Precision	100	100	100	100	100	96.887	100	100
Recall	89.744	100	100	100	98.162	100	99.605	98.295
F-Measure	94.595	100	100	100	99.072	98.419	99.802	99.14

Table 8: The performance of the proposed system using ResNet based features and DT classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	100	100	100	100	100
Specificity	100	100	100	100	100	100	100	100
Precision	100	100	100	100	100	100	100	100
Recall	100	100	100	100	100	100	100	100
F-Measure	100	100	100	100	100	100	100	100

Table 9: The performance of the proposed system using ResNet based features and K-NN classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	100	98.897	99.518	99.605	98.579
Specificity	100	100	100	99.934	99.788	99.606	100	99.849
Precision	100	100	100	99.401	98.897	98.804	100	99.427
Recall	100	100	100	100	99.897	99.518	99.605	98.579
F-Measure	100	100	100	99.699	99.394	99.159	99.802	99.001

Table 10: The performance of the proposed system using ResNet based features and SVM classifier

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	89.744	98.942	98.667	95.984	97.917	99.759	99.759	98.295
Specificity	100	99.959	99.891	99.956	99.433	99.054	99.837	100
Precision	100	98.942	98.886	99.583	97.084	97.183	99.079	100
Recall	89.744	98.942	98.667	95.984	97.917	99.759	99.209	98.295
F-Measure	94.595	98.942	98.776	97.75	97.498	98.454	99.144	99.14

Journal of Theoretical and Applied Information Technology
31st January 2020. Vol.98. No 02

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 11: The performance of the proposed system using ResNet based features and SVM classifier and cubic kernel function

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	100	100	100	99.598	98.897	99.759	100	98.863
Specificity	100	100	99.935	100	99.929	99.553	99.93	100
Precision	100	100	99.338	100	99.629	98.649	99.606	100
Recall	100	100	100	99.598	98.897	99.759	100	98.864
F-Measure	100	100	99.668	99.799	99.262	99.201	99.803	99.429

Table 12: The performance of the proposed system using ResNet based features and SVM classifier and Gaussian kernel function

Performance Metrics (%)	Age Group							
	A	B	C	D	E	F	G	H
Accuracy	89.743	100	100	100	98.162	100	99.605	98.011
Specificity	100	100	100	100	100	98.949	100	100
Precision	100	100	100	100	100	96.887	100	100
Recall	89.743	100	100	100	98.162	100	99.605	98.295
F-Measure	94.594	100	100	100	99.072	94.19	99.802	99.14

Table 13: A summary of the obtained results for the proposed system (%)

Method		Performance Measures				
		Accuracy	Specificity	Precision	Recall	F-measure
AlexNet	LD	99.3	99.9	99.479	99.575	99.526
	DT	99.5	99.917	99.617	99.636	99.626
	K-NN	99.3	99.916	99.63	99.748	99.689
	SVM-Quadratic	91.1	98.619	94.802	90.942	92.737
	SVM-Cubic	99.4	99.906	99.442	99.614	99.527
	SVM-Gaussian	99.2	99.869	99.611	98.226	98.879
ResNet	LD	N/A	N/A	N/A	N/A	N/A
	DT	100	100	100	100	100
	K-NN	99.3	99.897	99.566	99.699	99.632
	SVM-Quadratic	98.5	99.766	98.845	97.315	98.037
	SVM-Cubic	99.5	99.918	99.653	99.639	99.645
	SVM-Gaussian	99.2	99.869	99.611	98.226	98.879

5. CONCLUSION AND FUTURE WORK

Human age estimation became an important issue in our modern society. Manual age estimation methods have many drawbacks. Hence, there is a necessity to build efficient and effective automated human age estimation system. On the other side, transfer learning has proved its effectiveness in many machine learning and object recognition problem. In this paper, a transfer learning-based bone age assessment has been proposed using hand X-ray images. The proposed bone age assessment can classify the hand X-ray images into eight age groups. In the proposed system, the hand X-ray images are preprocessed, and the features are extracted using two well-known deep neural networks namely, AlexNet and ResNet. Finally, several classification models have been employed to perform the classification task including decision tree (DT), linear discriminant (LD), k-nearest neighbor (K-NN), and support vector machine (SVM). Based on the experimental results, the ResNet based features is better than the AlexNet based features. Also, the DT classifier is the best in terms of the different performance metrics compared to other classifiers. Hence, a promising performance has been achieved in the scope of automatic human age estimation. In the future, we intend to evaluate the proposed system using larger dataset and other classification models.

REFRENCES:

[1] Bakthula, Rajitha, and Suneeta Agarwal. "Automated human bone age assessment using image processing methods-survey." International Journal of Computer Applications 104, no. 13(2014): 33-42.

[2] Cattaneo, C., D. De Angelis, M. Ruspa, D. Gibelli, R. Cameriere, and M. Grandi. "How old am I? Age estimation in living adults: a case report." J Forensic Odontostomatol 26, no. 2 (2008): 39-43.

[3] Liversidge, Helen M., M. Christopher Dean, and Theya I. Molleson. "Increasing human tooth length between birth and 5.4 years." American Journal of Physical Anthropology 90, no. 3 (1993): 307-313.

[4] Cardoso, Hugo FV. "A test of the differential accuracy of the maxillary versus the mandibular dentition in age estimations of immature skeletal remains based on developing tooth length." Journal of forensic sciences 52, no. 2 (2007): 434-437.

[5] Yun, Jong-Il, Jeong-Yun Lee, Jin-Woo Chung, Hong-Seop Kho, and Young-Ku Kim. "Age estimation of Korean adults by occlusal tooth wear." Journal of forensic sciences 52, no. 3 (2007): 678-683.

[6] Cameriere, Roberto, Giuseppe Brogi, Luigi Ferrante, Dora Mirtella, Claudia Vultaggio, Mariano Cingolani, and Gino Fornaciari. "Reliability in age determination by pulp/tooth ratio in upper canines in skeletal remains." Journal of forensic sciences 51, no. 4 (2006): 861-864.

[7] Verma, Rajesh Kumar, Mukesh K. Goyal, and Shiv Kochar. "Age Assessment from Radiological Cranial Suture closure in Fourth to Seventh decades (A Jaipur Based Study)." Governing Council 2010-2012 32 (2010): 120-132.

[8] Vyas, P. C., and H. Sischer. "Age estimation by closure of suture of skull in individuals of 25-45 yrs of age of Jaipur area." PhD diss., Dissertation for MD (Forensic Medicine) University of Rajasthan, 1996.

[9] Atamturk, Derya, and Izzet Duyar. "Age-related factors in the relationship between foot measurements and living stature and body weight." Journal of forensic sciences 53, no. 6 (2008): 1296-1300.

[10] Hackman, Lucina, Catriona M. Davies, and Sue Black. "Age estimation using foot radiographs from a modern Scottish population." Journal of forensic sciences 58 (2013): S146-S150.

[11] Pyle, Sarah Idell, and Normand Louis Hoerr. A radiographic standard of reference for the growing knee. CC Thomas, 1955.

[12] Hackman, Lucina, and Sue Black. "Age estimation from radiographic images of the knee." Journal of forensic sciences 58, no. 3 (2013): 732-737.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[13] Iannaccone, G. "WW Greulich and SI Pyle: Radiographic atlas of skeletal development of the hand and wrist. I volume-atlante di 256 pagine. Stanford University Press, Stanford, California, 1959." Acta geneticae medicae et gemellologiae: twin research 8, no. 4 (1959): 513-513.

[14] Tanner, James Mourilyan, R. H. Whitehouse, N. Cameron, W. A. Marshall, M. J. R. Healy, and H. Goldstein. Assessment of skeletal maturity and prediction of adult height (TW2 method). Vol. 16. London: Academic Press, 1975.

[15] Tanner, James Mourilyan, R. H. Whitehouse, N. Cameron, W. A. Marshall, M. J. R. Healy, and H. Goldstein. Assessment of skeletal maturity and prediction of adult height (TW3 method). 3rd ed. London: W.B Saunders; 2001.

[16] Chumela, Wm Cameron, Alex F. Roche, and David Thissen. "The FELS method of assessing the skeletal maturity of the hand-wrist." American Journal of Human Biology 1, no. 2 (1989): 175-183.

[17] Michael, David J., and Alan C. Nelson. "HANDX: a model-based system for automatic segmentation of bones from digital hand radiographs." IEEE transactions on medical imaging 8, no. 1 (1989): 64-69.

[18] Pietka, Ewa, Michael F. McNitt-Gray, M. L. Kuo, and H. K. Huang. "Computer-assisted phalangeal analysis in skeletal age assessment." IEEE transactions on medical imaging 10, no. 4 (1991): 616-620.

[19] Garn, Stanley M., Keith P. Hertzog, Andrew K. Poznanski, and Jerrold M. Nagy. "Metacarpophalangeal length in the evaluation of skeletal malformation." Radiology 105, no. 2 (1972): 375-381.

[20] Tanner, James M., and Robert D. Gibbons. "A computerized image analysis system for estimating Tanner-Whitehouse 2 bone age." Hormone Research in Paediatrics 42, no. 6 (1994): 282-287.

[21] Gross, George W., John M. Boone, and Dorene M. Bishop. "Pediatric skeletal age: determination with neural networks." Radiology 195, no. 3 (1995): 689-695.

[22] Sato, Koshi, Kumi Ashizawa, Makoto Anzo, Fumio Otsuki, Shunichi Kaneko, Toshiaki Tanaka, Katsumi Tsukagoshi et al. "Setting up an automated system for evaluation of bone age." Endocrine journal 46, no. Suppl (1999): S97-S100.

[23] Hsieh, Chi-Wen, Tai-Lang Jong, and Chui-Mei Tiu. "Bone age estimation based on phalanx information with fuzzy constrain of carpals." Medical & biological engineering & computing 45, no. 3 (2007): 283-295.

[24] Thodberg, Hans Henrik, Sven Kreiborg, Anders Juul, and Karen Damgaard Pedersen. "The BoneXpert method for automated determination of skeletal maturity." IEEE transactions on medical imaging 28, no. 1 (2008): 52-66.

[25] Thodberg, Hans Henrik, Rick R. van Rijn, Oskar G. Jenni, and David D. Martin. "Automated determination of bone age from hand X-rays at the end of puberty and its applicability for age estimation." International journal of legal medicine 131, no. 3 (2017): 771-780.

[26] Lee, Hyunkwang, Shahein Tajmir, Jenny Lee, Maurice Zissen, Bethel Ayele Yeshiwas, Tarik K. Alkasab, Garry Choy, and Synho Do. "Fully automated deep learning system for bone age assessment." Journal of digital imaging 30, no. 4 (2017): 427-441.

[27] Spampinato, Concetto, Simone Palazzo, Daniela Giordano, Marco Aldinucci, and Rosalia Leonardi. "Deep learning for automated skeletal bone age assessment in X-ray images." Medical image analysis 36 (2017): 41-51.

[28] Chu, Meicheng, Bo Liu, Fugen Zhou, Xiangzhi Bai, and Bin Guo. "Bone Age Assessment Based on Two-Stage Deep Neural Networks." In 2018 Digital Image

Computing: Techniques and Applications (DICTA), pp. 1-6. IEEE, 2018.

[29] Kaur, Amandeep, and Kulwinder Singh Mann. "Segmenting Bone Parts for Bone Age Assessment using Point Distribution Model and Contour Modelling." In Journal of Physics: Conference Series, vol. 933, no. 1, p. 012004. IOP Publishing, 2018.

[30] Mutasa, Simukayi, Peter D. Chang, Carrie Ruzal-Shapiro, and Rama Ayyala. "MABAL: a novel deep-learning architecture for machine-assisted bone age labeling." Journal of digital imaging 31, no. 4 (2018): 513-519.

[31] Tajmir, Shahein H., Hyunkwang Lee, Randheer Shailam, Heather I. Gale, Jie C. Nguyen, Sjirk J. Westra, Ruth Lim, Sehyo Yune, Michael S. Gee, and Synho Do. "Artificial intelligence-assisted interpretation of bone age radiographs improves accuracy and decreases variability." Skeletal Radiology 48, no. 2 (2019): 275-283.

[32] Yu, Wei, Jing Chang, Cheng Yang, Limin Zhang, Han Shen, Yongquan Xia, and Jin Sha. "Automatic classification of leukocytes using deep neural network." In 2017 IEEE 12th International Conference on ASIC (ASICON), pp. 1041-1044. IEEE, 2017.

[33] Thanh, T. T. P., Caleb Vununu, Sukhrob Atoev, Suk-Hwan Lee, and Ki-Ryong Kwon. "Leukemia blood cell image classification using convolutional neural network." International Journal of Computer Theory and Engineering 10, no. 2 (2018): 54-58.

[34] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097-1105. 2012.

[35] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.

[36] Transfer learning, https://medium.com/@14prakash/transfer-learning-using-keras-d804b2e04ef8, Last Access: 15-10-2019.

[37] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Identity mappings in deep residual networks." In European conference on computer vision, pp. 630-645. Springer, Cham, 2016.

[38] Quinlan, J. Ross. "Induction of decision trees." Machine learning 1, no. 1 (1986): 81-106.

[39] Cover, Thomas M., and Peter E. Hart. "Nearest neighbor pattern classification." IEEE transactions on information theory13, no. 1 (1967): 21-27.

[40] Zhao, W., Rama Chellappa, and Nagaraj Nandhakumar. "Empirical performance analysis of linear discriminant classifiers." In Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231), pp. 164-169. IEEE, 1998.

[41] Vapnik, V., and Vlamimir Vapnik. "Statistical learning theory Wiley." New York (1998): 156-160.

[42] https://www.kaggle.com/kmader/rsna-boneage?fbclid=IwAR28eIZr97SQpE7s3yq2Z4YObfUeOFlNt SDKC1xY1TICClROT1J2XdY-Lg, Last Access: 15-10-2019.

[43] https://www.mathworks.com/matlabcentral/fileexchange/33170-multi-class-support-vector-machine, Last Access: 15-10-2019.