TW3-Based Fully Automated Bone Age Assessment System Using Deep Neural Networks

SUNG JOON SON[1], **YOUNGMIN SONG**[2], **NAMGI KIM**[2], **YOUNGHAE DO**[3], **NOJUN KWAK**[1], **(Member, IEEE), MU SOOK LEE**[4], **AND BYOUNG-DAI LEE**[2,5]

[1]Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, South Korea
[2]Department of Computer Science, Kyonggi University, Suwon 16227, South Korea
[3]Department of Mathematics, Kyungpook National University, Daegu 41566, South Korea
[4]Human Medical Imaging Center, Seoul 06524, South Korea
[5]BLEE Co., Suwon 16227, South Korea

Corresponding author: Byoung-Dai Lee (blee@kgu.ac.kr)

ABSTRACT Deep learning technology has rapidly evolved in recent years. Bone age assessment (BAA) is a typical object detection and classification problem that would benefit from deep learning. Convolutional neural networks (CNNs) and their variants are hence increasingly used for automating BAA, and they have shown promising results. In this paper, we propose a complete end-to-end BAA system to automate the entire process of the Tanner–Whitehouse 3 method, starting from localization of the epiphysis–metaphysis growth regions within 13 different bones and ending with estimation of the corresponding BA. Specific modifications to the CNNs and other stages are proposed to improve results. In addition, an annotated database of 3300 X-ray images is built to train and evaluate the system. The experimental results show that the average top-1 and top-2 prediction accuracies for skeletal bone maturity levels for 13 regions of interest are 79.6% and 97.2%, respectively. The mean absolute error and root mean squared error in age prediction are 0.46 years and 0.62 years, respectively, and accuracy within one year of the ground truth of 97.6% is achieved. The proposed system is shown to outperform a commercially available Greulich–Pyle-based system, demonstrating the potential for practical clinical use.

INDEX TERMS Bone age assessment, deep learning, GP, TW3.

I. INTRODUCTION

The growth of children is affected not only by genetic, hormonal, and nutritional factors, but also by diseases and psychosocial elements. Therefore, significant deviations from normal growth may suggest signs of various diseases, such as endocrine disorders, pediatric syndromes, and even genetic disorders [1], [2]. Of the different growth assessment methods, those based on skeletal maturity or bone age (BA) are the most commonly used in a clinical setting owing to their simplicity, minimal radiation exposure, and availability of multiple ossification centers [3].

If your paper is intended for a conference, please contact your conference editor concerning acceptable word

The associate editor coordinating the review of this manuscript and approving it for publication was Naveed Akhtar.

processor formats for your particular conference. Although there is no general standard method for bone age assessment (BAA), the Greulich–Pyle (GP) method [4] and the Tanner–Whitehouse (TW) method [5] are the most typically employed in clinical practice. The GP method is based on a hand atlas, which consists of a series of template X-ray images covering the growth stages of children with varying bone maturity levels. A patient's X-ray image is then compared against samples in the template series, and the closest matching template is chosen as the equivalent BA of the patient. This approach is straightforward and can be quickly performed. However, the GP method is marked by inter- and intra-observer variability. Moreover, it is difficult to accurately assess bones with large variations, and the resulting BAs are coarse-grained because the template series is arranged in intervals ranging from six months to one year.

Instead of using all the bones of the left hand, as in the GP method, the TW method evaluates the maturity levels of specific bones of the hand and wrist, which comprise the region of interest (ROI). Predefined skeletal maturity scores are assigned to individual ROIs in accordance with their maturity levels. These scores are then summed to compute the total maturity score. This score is finally converted into a BA using a correlation matrix for the maturity scores and BAs. The TW method has been revised several times. Currently, the TW3 analyzes the radius, ulna, and short bones (RUS) as the ROIs (see Fig. 1). Compared to the GP method, the TW3 method is comparatively more complex and requires more time; however, it is more accurate and reproducible [6] because the correlation matrix is organized using units of 0.1 years. In addition, because the maturity of each bone is evaluated and scored independently, the TW3 method is robust to mispredictions in some bones, leading to a more objective evaluation.
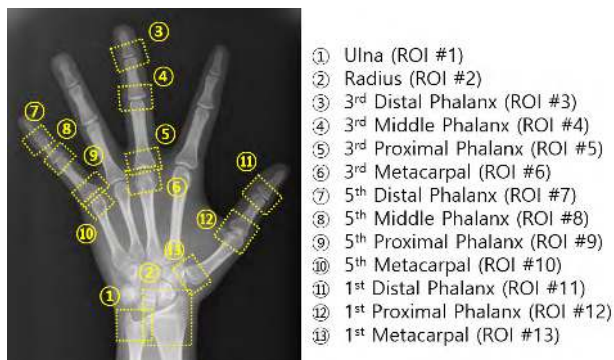

① Ulna (ROI #1)
② Radius (ROI #2)
③ 3rd Distal Phalanx (ROI #3)
④ 3rd Middle Phalanx (ROI #4)
⑤ 3rd Proximal Phalanx (ROI #5)
⑥ 3rd Metacarpal (ROI #6)
⑦ 5th Distal Phalanx (ROI #7)
⑧ 5th Middle Phalanx (ROI #8)
⑨ 5th Proximal Phalanx (ROI #9)
⑩ 5th Metacarpal (ROI #10)
⑪ 1st Distal Phalanx (ROI #11)
⑫ 1st Proximal Phalanx (ROI #12)
⑬ 1st Metacarpal (ROI #13)

FIGURE 1. ROIs used in the TW3 method.

Deep learning technologies that are based on artificial neural networks have rapidly evolved in recent years, and their application areas have also expanded. In particular, deep learning techniques have permeated the entire field of medical image analysis [7] and have achieved state-of-the-art results in some medical imaging applications, such as the segmentation of brain lesions [8], leak detection in airway tree segmentation [9], diabetic retinopathy classification [10], and breast cancer metastasis detection [11]. The BAA task is a typical object detection and classification problem that could benefit from deep learning. In this approach, for a given input (e.g., a left hand and wrist X-ray image of a patient), a corresponding class (e.g., a class corresponding to a BA) is predicted. Convolutional neural networks (CNNs) and their variants are thus increasingly used for automating BAA, and they have shown promising results [3], [12], [13], [17].

It is recognized that the TW3 method outperforms the GP method in terms of assessment accuracy and objectivity. Nevertheless, to the best of our knowledge, none of the existing systems are truly TW3-based, fully automated end-to-end BAA systems. This is largely attributed to the following issues. First, it is not easy to accurately locate all of the ROIs within an X-ray image of the left hand and wrist.

In particular, the maturity indicators of individual bones—such as the epiphysis–metaphysis length ratio, existence of the styloid process, and onset of epiphyseal–diaphyseal fusion—can be found only in the epiphysis—metaphysis growth regions of the bones. Consequently, failure to acquire the precise location information of a growth region of any single ROI can degrade assessment accuracy. Although segmentation algorithms based on conventional image processing techniques have been proposed to detect the contours of bones, they are not adequately robust to handle bones of varying sizes, shapes, morphologies, and densities [14], [15]. Second, the lack of large-scale labeled training datasets is another critical obstacle to the development of TW3-based fully automated BAA systems. Preparation of such datasets requires significant time and expenses because, for a given X-ray image of the left hand and wrist, the radiologist readings of the corresponding BA and maturity levels of individual ROIs are required for deep learning algorithms to learn the TW3 method.

In this paper, we therefore propose a novel TW3-based fully automated BAA system using deep neural networks. The proposed system is believed to be the first complete end-to-end system to automate the entire process of the TW3 method, starting from localization of the epiphysis–metaphysis growth regions of 13 ROIs and ending with estimation of the corresponding BA. To that end, we also built a large training dataset, which contains approximately 3,300 X-ray images and the corresponding radiologist readings using the TW3 method. In addition, we developed techniques based on CNNs and region-based CNNs (R-CNNs) for the classification of the maturity level of an ROI and the localization of the epiphysis–metaphysis growth regions of ROIs, respectively. To assess the effectiveness of the proposed BAA system, we performed an evaluation using randomly selected test X-ray images and we compared it to a GP-based commercial product. The remainder of this paper is organized as follows. Section II summarizes related work. Section III details the proposed system. Section IV presents the experimental results, and Section V summarizes the research and directions for future work.

II. RELATED WORK

The objective of this study is automatic TW3-based BAA using deep neural networks. Thus, our review of related research primarily focuses on deep-learning-based approaches. Comprehensive surveys on computerized methods for BAA are provided in [14] and [15].

In [3], the unique characteristics of X-ray images of the left hand and wrist are considered in a GP-based CNN called BoNet, which consists of five convolutional layers, one deformation layer, and one 2048-fully connected layer followed by a single output neuron. To investigate the performance of BoNet, extensive experiments varying the configuration parameters were performed. In general, deeper networks perform better. Nevertheless, it was found that using as many convolutional layers as possible does not necessarily mean a

high prediction accuracy, and the best layer depth is strongly domain-dependent.

In [12], a GP-based fully automated deep learning platform was proposed for BAA. The system is composed of two main components: a preprocessing engine and a classifier. The preprocessing engine divides the entire 512×512 X-ray image into sample patches (e.g., 32×32 image patches) and the class of each patch is determined by a CNN. Using the image patch classification results, a mask for the hand and wrist is created and enhanced by image processing. For a preprocessed X-ray image, the classifier estimates BA using a CNN consisting of a fine-tuned pretrained GoogLeNet model [19]. Although the proposed system is claimed to be effective and robust in assessing the BAs of various X-ray images, it provides integer-based coarse-grained estimations, which is an intrinsic problem of the GP method.

The approach presented in [13] is similar to that of [12] in that it employs existing models (e.g., GoogLeNet and VGGNet [20]) for weight initialization and fine-tuning is performed across all layers to develop a model for predicting BAs. According to their empirical analysis, outliers in the BA predictions can occur because the cross-entropy loss considers only the correctness of the predictions, regardless of the degree of incorrectness. Hence, an L2-based loss function was proposed to measure the difference between the predicted BA and ground truth. Although the overall prediction accuracy of the proposed solution is not remarkable, this approach can achieve competent results that are close to a radiologist's readings.

Introduced in 2009 and based on conventional image processing techniques, BoneXpert [16] is commercially available software for automated BAA. BoneXpert consists of three layers. Layer A reconstructs bone borders using an active appearance model. Layer B derives an intrinsic BA based on the chronological age and appearance of the bones. Finally, layer C transforms the computed intrinsic BA into either a GP BA or TW2 BA. Although BoneXpert was clinically approved and validated for various ethnicities, it has several critical limitations. To apply it, the relationship between BAs and chronological ages must be studied for target ethnicities and provided as input [21]. Moreover, it is not robust to abnormal X-ray images. For instance, images with a relatively large amount of noise or incorrectly posed bones are rejected. It was reported that BoneXpert rejected approximately 235 individual bones out of 5,161 (4.5%) [12].

VUNOMed-Bone Age [17] is another commercialized BAA software that uses a CNN-based deep learning engine to determine a patient's BA based on GP. Unlike most other BAA systems, which estimate BAs, the software displays the three most likely age-labeled reference images given the patient's radiograph. According to [17], the software increases diagnostic accuracy by 8% and reduces diagnosis time by 40%.

Because of their modular structure, TW methods have been used in many previous studies on BAA automation. The accurate extraction of ROIs is a fundamental process in TW methods for which various segmentation-based techniques have been proposed. However, the majority of these methods are limited because manual ROI localization is required or X-ray images that deviate from the normal form cannot be processed [22]–[26].

In short, previous deep learning-based BAA systems all employ the GP method. Thus, they suffer from the lack of measurement repeatability and the systematic errors that are inherent to the GP method [27]. In contrast, this study presents a TW3-based, complete end-to-end BAA system using deep neural networks.

III. METHODLOGY

As shown in Fig. 2, the proposed BAA system consists of three steps. The first step is to extract sufficiently large areas of the wrist, thumb, middle finger, and little finger, including the actual ROIs that are used in the TW3 method. We call these areas the bounding ROIs (bROIs). Figure 3 illustrates examples of the four bROIs. The actual ROIs contained in the individual bROIs are the following:

- bROI of the wrist: the radius and ulna;
- bROI of the thumb: the first distal phalanx, first proximal phalanx, and first metacarpal;
- bROI of the middle finger: the third distal phalanx, third middle phalanx, third proximal phalanx, and third metacarpal;
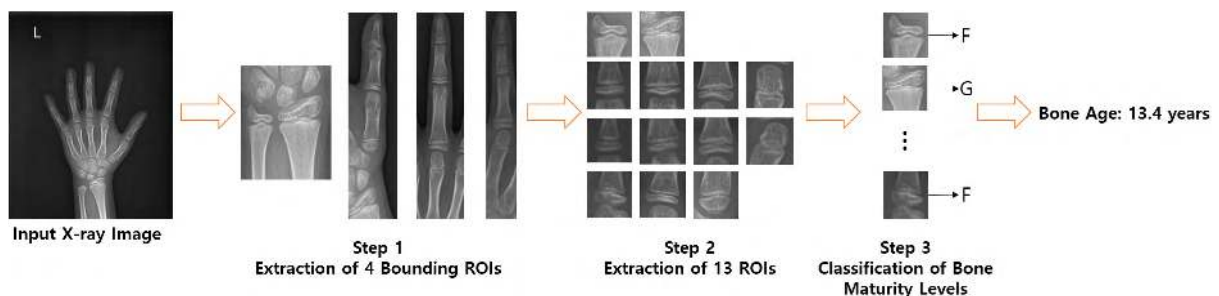- bROI of the little finger: the fifth middle phalanx, fifth proximal phalanx, and fifth metacarpal.

FIGURE 2. Proposed automated bone age assessment process.

FIGURE 3. Examples of bounding ROIs.

The second step extracts the 13 actual ROIs from the bROIs. The proposed ROI extraction technique is a hybrid in that the bROI extraction is conventional image processing, whereas the extraction of the actual ROIs uses deep learning techniques. The third step classifies the levels of the skeletal maturity of the individual ROIs and converts them into scores. Finally, the bone age is predicted using these scores and the correlation matrix. We next present the detailed process of each step.

A. STEP1: EXTRACTION OF bROIs

The first process to extract bROIs is to rotate the original image to make the wrist vertical. Then, different algorithms are applied depending on the location of the bROIs,

i.e., the wrist and fingers. Figure 4 shows the intermediate outputs of the bROI extraction processes. The blue points are used to set boundaries, and the red box represents the bROI.

For the wrist rotation process, a given input X-ray image is first converted into a grayscale image and binarized. The binarized image, in turn, is modified so that all pixels apart from the border pixels are set to black to extract the outline of the hand. Then, the coordinates of the white pixels of each row are stored. This is performed for an arbitrary number of rows from the bottom to the top of the image. Using these pixel points, the midpoints of the wrist are computed and then the straight line proximate to these points is computed. This line can be obtained using various methods, such as regression analysis or random sample consensus. Finally, the tilt of the straight line is computed and used to rotate the image to make the wrist vertical.

Once the wrist has been rotated, the original and rotated coordinates of the endpoints of the wrist are compared. The higher y-coordinate forms the lower boundary line of the bROI of the wrist. The center point of the hand is computed using the hand outline image. Various methods, including moments, can be applied to calculate the center point. The calculated center point of the hand forms the upper boundary of the bROI of the wrist. To obtain the left and right boundaries of the bROI, the procedure used to compute the straight line proximate to the midpoints of the wrist is applied. For the left boundary, for example, pixel coordinates comprising the left outline of the wrist are first collected and the straight line perpendicular to the x-axis is approximated using
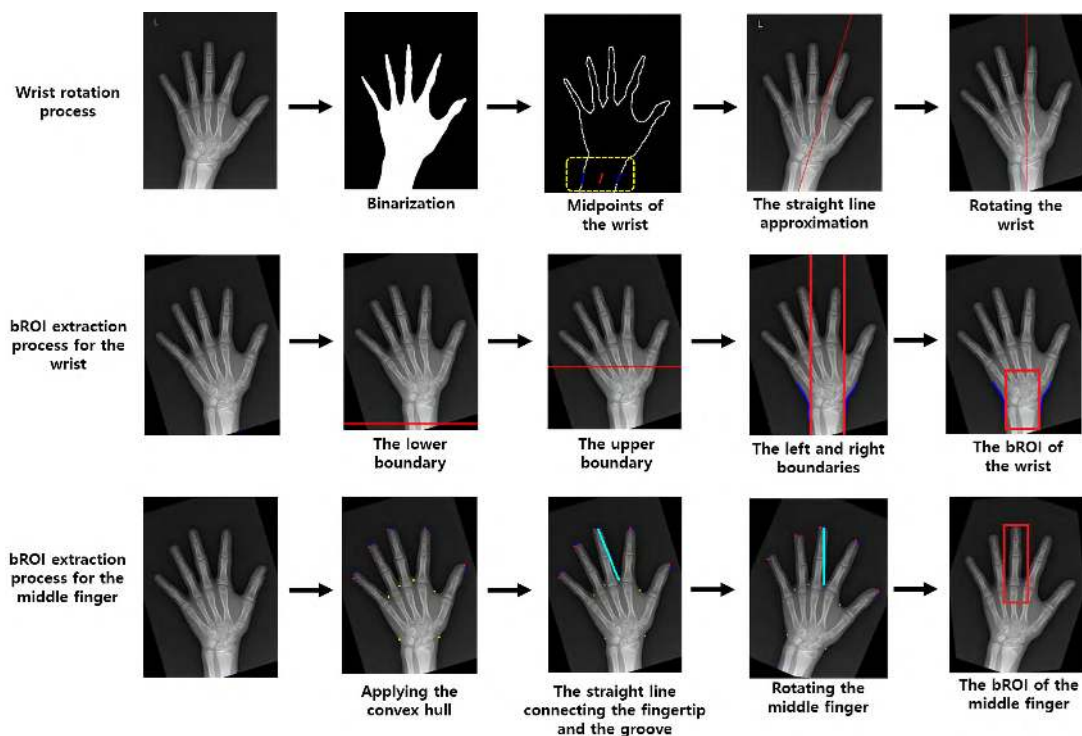

FIGURE 4. Intermediate outputs of the bROI extraction processes.

regression analysis. The final bROI of the wrist is extracted from the rotated image using the upper, lower, left, and right boundaries.

To extract bROIs from the fingers, a convex hull is calculated for the rotated X-ray image to determine finger points. Then, for each bROI, a straight line connecting the tip of the corresponding finger and either the left or right groove is identified. Then, the image is rotated to make the straight line vertical. Finally, the upper and lower boundaries of the bROI are set using the y-coordinate of the fingertip and the y-coordinate of the center of the hand. The left and right boundaries are set using the x-coordinates of the grooves on both sides.

B. STEP2: EXTRACTION OF ACTUAL ROIs

To extract the actual ROIs from a bROI, we apply the faster region-based CNN (Faster R-CNN) [28]. Figure 5 illustrates the Faster R-CNN architecture used for this task. The first convolutional layer extracts the feature map of the input image that the region proposal network (RPN) uses to identify ROI candidates using a sliding window. As the sliding window sequentially shifts in the feature map, nine anchor boxes that differ in scale (128, 256, and 512) and aspect ratio (1:1, 1:2, and 2:1) based on its center are determined. Then, fixed-size feature vectors are extracted from the candidates using ROI pooling and are input to the fully connected layer. Finally, the sizes and locations of the final regions, as well as the prediction scores, are determined through two fully connected layers. We fine-tuned the Faster R-CNN while varying several hyperparameters. For instance, the overlap thresholds as well as the non-maximum suppression thresholds for the individual ROIs used in the RPN were selected by a grid search on the validation dataset.
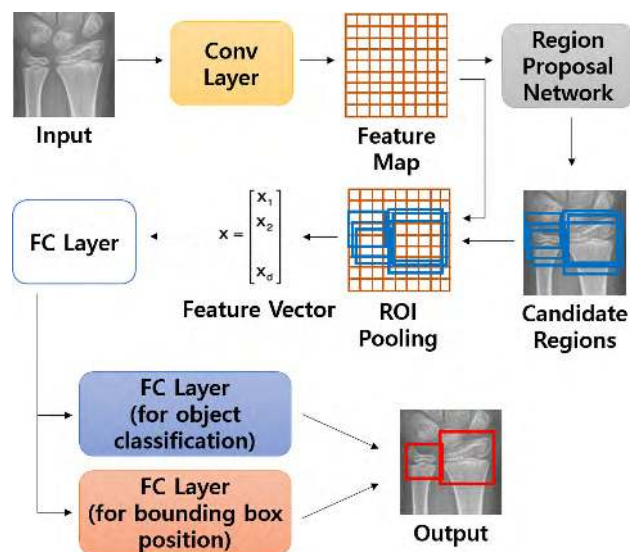

FIGURE 5. Workflow of faster R-CNN (Conv layer: convolutional layers, FC layer: fully connected layers).

Note that each individual bROIs has its own characteristics, such as the number and relative locations of actual ROIs within the bROIs. Therefore, we deployed four different models using Faster R-CNN, each of which is responsible for extracting the actual ROIs from the corresponding bROIs.

C. STEP3: BAA

The operations in the final step of the proposed BAA system are to evaluate the skeletal maturity levels of individual ROIs, determine their maturity scores, find the total RUS score from the RUS score table, and predict the BA using a correlation matrix. In particular, the automated classification of the skeletal maturity levels is a critical and challenging task. To tackle this problem and verify the feasibility of our approach, we developed the VGGNet-BA CNN. As the names of the proposed CNN suggests, we used VGGNet as the baseline network architecture and modified it to be able to learn local discriminative features from a relatively small number of small grayscale ROI images while reducing overfitting. Although it is not state-of-the-art, the reason that we decided to customize and enhance VGGNet for the task of classification of skeletal maturity levels is that VGGNet is conceptually simple and relatively easy to implement and train. In fact, it has been used as a baseline architecture in many research projects for the same reason. Note, however, that we do not claim that the proposed deep neural network architecture is an optimal solution for the classification task. Therefore, one of our future work is to create a deeper network by enabling additional convolutional layers equipped with state-of-the-art building blocks, such as residual blocks in ResNet [18] and dense blocks in DenseNet [37], to improve the feature representation ability of the network.
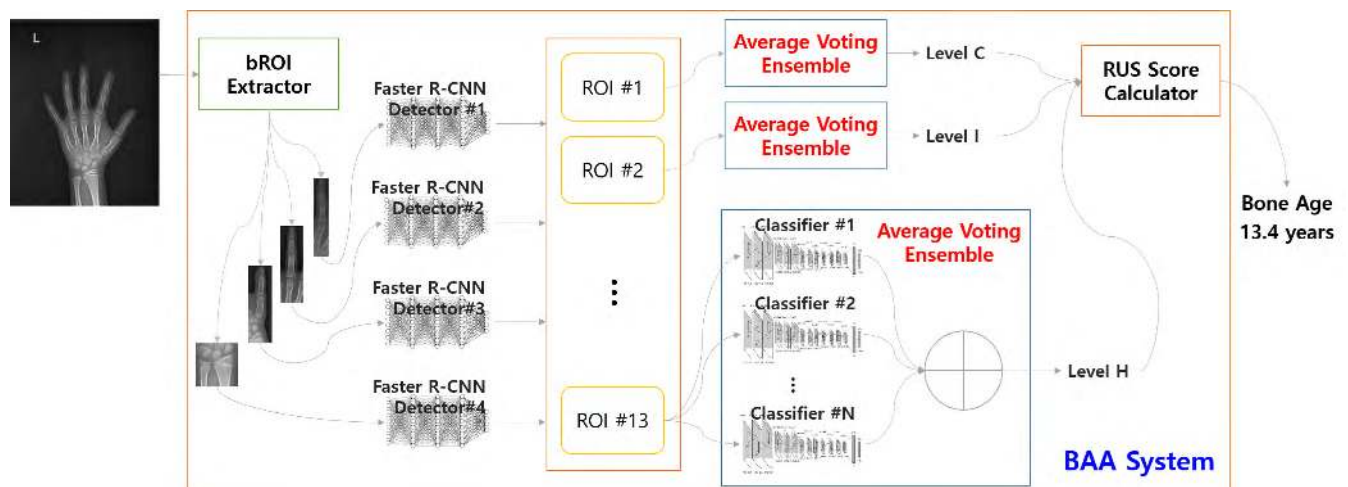
Table 1 shows the proposed VGGNet-BA configuration. The general architecture of VGGNet-BA is similar to those of VGGNet configurations in [20] in that a stack of convolutional layers is followed by two fully connected layers. However, key differences exist in the numbers of weighted layer, the number of channels for individual convolutional layers as well as the number and structure of the fully connected layers. As a result, the number of trainable parameters is significantly reduced and is approximately 750 times smaller than existing VGGNet configurations.

Another important difference is the input image size, which is changed from 224×224 to 112×112. In general, ROI images are smaller than the natural images used in image classification tasks (e.g., the ImageNet competition), and their sizes vary according to ROI locations and bone maturity levels. For instance, the size of a first distal phalanx image corresponding to level A is approximately 50×50, whereas the size of an ROI image containing the radius of level I is 220×220. However, a significant number of ROI image samples are approximately 110×110 pixels in size. Therefore, the proposed CNN accepts fixed-sized 112×112 input images to reduce image transformation due to image resizing.

In addition, we employ a simple average voting ensemble model in our system. That is, for a given input ROI image,

TABLE 1. The configuration of VGGNet-BA (FC: fully connected layer, s: Stride).

	LAYER 1	LAYER 2	LAYER 3	LAYER 4	LAYER 5	LAYER 6
VGGNET-BA	$\begin{bmatrix} conv3\times3, 8 \\ ReLU \\ conv3\times3, 8 \\ ReLU \\ MaxPool \\ (2\times2, s=2) \end{bmatrix}$	$\begin{bmatrix} conv3\times3, 16 \\ ReLU \\ conv3\times3, 16 \\ ReLU \\ MaxPool \\ (2\times2, s=2) \end{bmatrix}$	$\begin{bmatrix} conv3\times3, 32 \\ ReLU \\ conv3\times3, 32 \\ ReLU \\ MaxPool \\ (2\times2, s=2) \end{bmatrix}$	$\begin{bmatrix} conv3\times3, 64 \\ ReLU \\ conv3\times3, 64 \\ ReLU \\ MaxPool \\ (2\times2, s=2) \end{bmatrix}$	$\begin{bmatrix} FC-512d \\ ReLU \\ Dropout \end{bmatrix}$	$\begin{bmatrix} FC-8(9)d \\ SoftMax \end{bmatrix}$
# OF WEIGHTS (208,392/208,904)	648	3,456	13,824	55,296	131,072	4,096(4,608)

FIGURE 6. System architecture of the proposed deep-learning based BAA system.

an ensemble of independent CNN models predict the skeletal maturity levels and probabilities. The final prediction is generated by averaging the predictions of the individual models. Each individual ROI has unique features indicating the bone maturity level. Therefore, the proposed BAA system consists of 13 average voting ensembles, each of which is responsible for classifying the bone maturity level of the corresponding ROI. Figure 6 depicts the overall system architecture of the proposed BAA system.

IV. EXPERIMENTAL RESULTS
In this section, we describe the dataset and experimental results obtained from the proposed BAA system.

A. DATASET
Because no publicly available sample data exist for TW3-based machine learning, we created our own dataset consisting of 3,344 left hand and wrist X-ray images of Korean children under the age of 18. Each image has a different size and resolution. Each individual X-ray image is associated with five auxiliary files: four bROI images and an annotation file. The annotation file contains the coordinates of the 13 actual ROIs within the corresponding bROIs, the skeletal maturity levels of individual ROIs, and the estimated BA. The skeletal maturity levels of individual ROIs

TABLE 2. Distribution of dataset images by skeletal maturity levels.

ROI No.	LEVEL									TOTAL
	A	B	C	D	E	F	G	H	I	
1	1,035	167	239	141	280	986	342	154	✕	3,344
2	17	13	61	120	104	1,511	1,099	292	127	3,344
3	26	16	23	715	321	1,809	199	42	193	3,344
4	18	14	20	370	1,358	947	427	54	136	3,344
5	5	3	7	302	1,627	867	308	75	150	3,344
6	19	20	41	189	669	1,487	669	110	140	3,344
7	92	40	49	658	953	1,164	151	41	196	3,344
8	70	49	60	425	1,399	847	220	107	167	3,344
9	24	15	22	373	1,539	791	358	62	160	3,344
10	41	37	36	271	1,202	691	833	105	128	3,344
11	20	4	12	650	691	1,265	424	76	202	3,344
12	68	67	37	197	1,282	1,122	371	54	146	3,344
13	91	77	94	993	617	806	423	77	166	3,344
TOTAL	1,526	522	701	5,404	12,042	14,293	5,824	1,249	1,911	

were agreed upon by two professional radiology doctors to ensure the BAA accuracy. The distribution of X-ray images in the dataset according to the skeletal maturity level of individual ROIs is shown in Table 2. For training the deep neural networks, we used all the samples in the dataset except those in the test datasets presented in Table 4. In particular,

to train the Faster R-CNN for the actual ROI extraction, we randomly selected 20% of the sample X-ray images for use as a validation dataset, whereas the remaining 80% of the sample images were used as the training dataset. Similarly, to train the proposed CNN for the classification of the skeletal maturity levels, we divided the sample ROI images of individual skeletal maturity levels in such a way that 80% were used for model training and the remaining 20% for the validation.

B. ROI EXTRACTION DETAILS

We implemented the bROI extraction module using an open-source computer vision library (OpenCV 2.4.13.3) [29]. For the actual ROI extraction from the bROIs, we used the official Faster R-CNN code [30] and trained it from the beginning on our own X-ray image dataset. All four Faster R-CNN models used VGGNet-16 to train the RPN as well as the object detection network. The training was carried out using the stochastic gradient descent algorithm with momentum. The mini-batch size was set to 256 and the momentum was set to 0.9. The training was regularized by weight decay, which was set to 0.0005. The learning rate was initially set to 10^{-3}, and was decreased in steps by a factor of 10 every 100 iterations. The epoch size was set to 30. All models shared the above-mentioned hyperparameters.

C. TRAINING FOR CLASSIFICATION DETAILS

During training of the proposed CNN for classification of skeletal maturity levels, we artificially increased the size of the training dataset through simple data augmentation to reduce overfitting and to achieve high classification accuracy. The process for training data preparation was as follows. For a given X-ray image, its 13 source ROI images were generated by extracting rectangular areas specified by the ROI coordinate information in the corresponding annotation file. Each source ROI image was then resized to 134×134 and rotated by $+3°$ and $-3°$, respectively. These three intermediate ROI images (e.g., the resized source ROI image and two rotated images from the resized source ROI image) were cropped at the center by 119×119, followed by four 112×112 crops from the four corners of the images. As a result, 12 synthetic images were generated from a single source ROI image. The data augmentation improved the overall prediction accuracy by 1%–2%.

Examples of source ROI images resized to be fed into the CNN are presented in Fig. 7. The figure clearly shows that, as growth progresses, the shape, size, and texture of individual ROIs change accordingly. In addition, the ROIs of immediate neighboring levels, particularly those belonging to levels E to H, are quite similar. It thus requires sophisticated training to determine exactly to which level an ROI belongs.

The current implementation of the average voting ensemble consists of a set of independent VGGNet-BA models, each of which was initialized with different random weights drawn from the HE uniform initializer [31]. To prepare training data for individual component models, we divided the entire training dataset for a given ROI into N disjoint subsets,

where N represents the number of component models for the ROI. In addition, to avoid bias of the component models to certain levels, each subset was carefully prepared to ensure an even distribution of individual levels.

The hyperparameters used to train the component models of VGGNet-BA were the following. The Adam optimizer was used with a mini-batch size of 256 and a momentum of 0.9. The base learning rate was set to 10^{-3} and then decreased by a factor of 10 when the validation set accuracy ceased improving. The dropout rate was set to 0.5 for hidden units. The epoch size was set to 30. However, we applied early stopping to avoid overfitting. According to a recent study [32], using the same number of independent component classifiers as class labels gives the highest accuracy. Therefore, individual average voting ensembles of the proposed BAA system had nine components models except for the one for the ulna, which had eight components models.

Owing to a lack of training samples, some maturity levels in the ROIs, particularly levels A to C, were not included in the deep learning for classification, except for the ulna. As shown in Fig. 7, the early stages of bone growth (e.g., levels A to D) are relatively evident for evaluation. Therefore, we believe that acquisition of sufficient sample data for these levels could resolve the problem. In contrast, for levels E to G of the ulna and levels E to H of the other ROIs, the neighboring levels have a similar appearance. It is hence a more difficult and time-consuming task to distinguish those levels accurately [33]. Therefore, we believe that it is reasonable to evaluate performance by measuring how accurately a BAA system can classify these levels automatically. Table 3 shows the skeletal maturity levels of ROIs used in the classification.

TABLE 3. Skeletal maturity levels used for classification.

ROI No.	Levels used in classification
1	A, B, C, D, E, F, G, H
2	F, G, H, I
3–13	D, E, F, G, H, I

Another issue with the current dataset is class imbalance. For instance, levels E and F contain more samples than levels H and I. To address the class imbalance, we modified the loss function of VGGNet-BA to apply class-weights during the training, as shown in (1). The class-weights scale the calculated loss for each observation by the appropriate class weight such that more significance is placed on the losses associated with the minority classes [34].

$$\text{Loss} = -\sum_{k=1}^{n} w_k t_k log\left(\hat{y}_k\right) \qquad \text{Eq.(1)}$$

where $\hat{y}_k$, t_k, and w_k are one-hot encoded vectors that represent the predicted skeletal maturity level distribution, corresponding ground truth, and class weights, respectively. For the experiments, we used the scikit –learn package [35] to compute the class-weights.
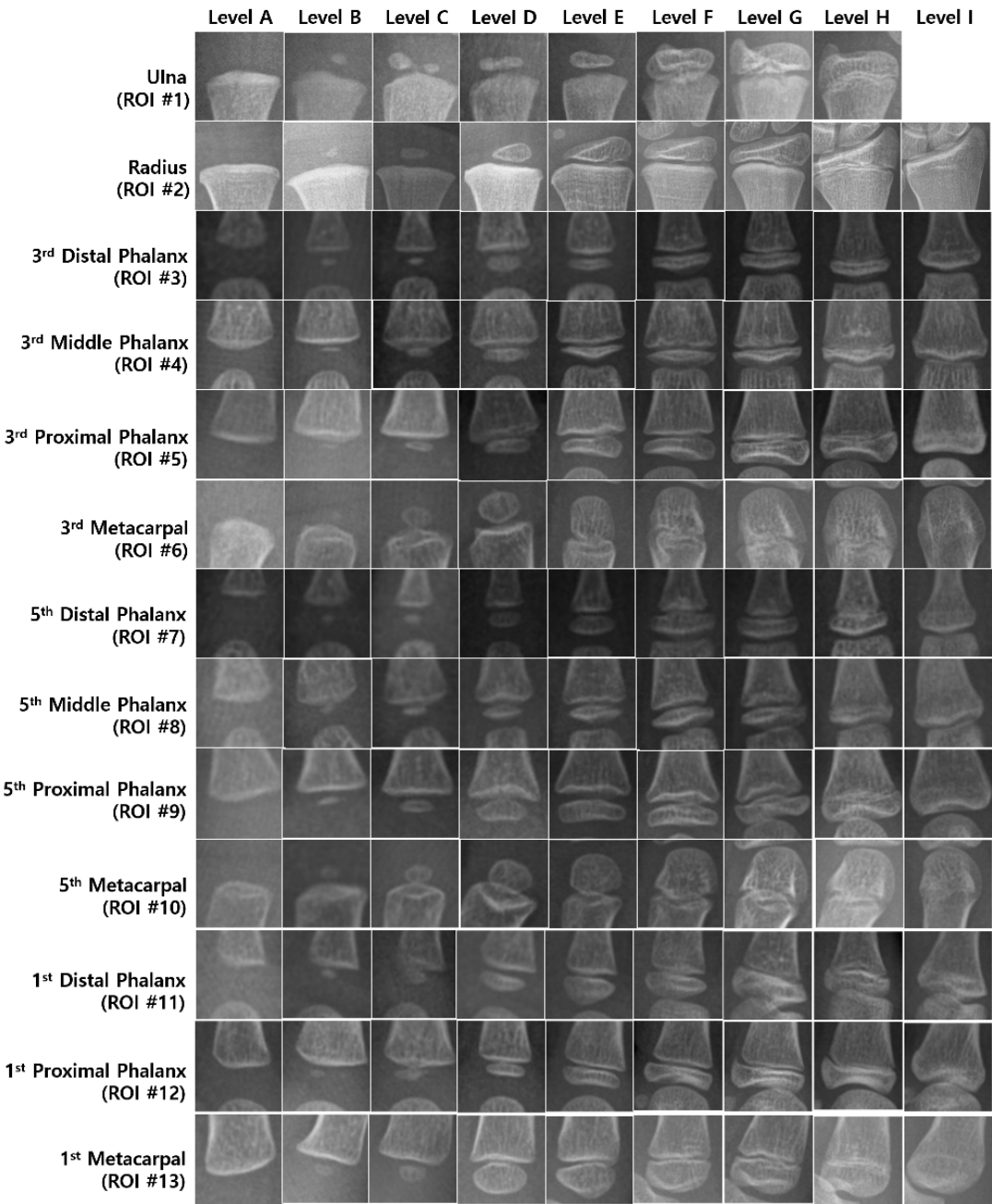
FIGURE 7. Example images of skeletal maturity levels of individual ROIs.

D. EVALUATION

We randomly generated two test datasets, as shown in Table 4. Test dataset 1 was used to evaluate the ROI extraction accuracy and BA prediction accuracy of the proposed deep-learning based BAA system. Test dataset 2 is a subset of dataset 1 and was used to compare the proposed method with a commercialized BAA system that is based on the GP method (hereafter, referred to as X-BAA Service). The number of test samples in dataset 2 was limited by the number of credits that we purchased for X-BAA Service. Along with the total number of test samples, another difference between the two datasets was the exclusion of samples belonging to subjects aged 15 to 18 years. In the TW3 method, if the total maturity score exceeds 1,000, all of patients are considered to be of the same BA. However, GP enables distinction up to 18 years of age. Therefore, for an accurate and fair evaluation, we decided not to include samples for subjects aged 15 to 18 years in dataset 2.

To improve the validity of the proposed BAA system, we applied a 10-fold cross validation on the test datasets: for

TABLE 4. Distribution of test datasets images by chronical ages.

AGE (YEARS)	NUMBER OF SAMPLES	
	TEST DATASET 1	TEST DATASET 2
2–4	20	11
5–6	80	34
7–8	220	65
9–10	100	27
11–12	50	16
13–14	20	7
15–18	10	
TOTAL	500	160

each fold we measured the corresponding performance metrics, and the final result was the average of the 10 evaluations.

1) ROI EXTRACTION ACCURACY

To assess the effectiveness of our approach based on bROI, we used the official Faster R-CNN code and trained it with X-ray images annotated with 13 ROI coordinates within the images. We compared the success rate of ROI extraction of the two methods. For a given X-ray image, when the following two conditions were met, the extraction was deemed a success.

① All 13 ROIs were automatically extracted.
② The skeletal maturity levels of all ROIs could be predicted in the subsequent BAA step.

Table 5 shows the success rate of ROI extraction of the two methods. The rejection rate of our approach is approximately 1.6% and is notably lower than that of the comparative method.

TABLE 5. Success rate of ROI extraction.

ROI EXTRACTION METHOD	SUCCESS RATE
with bROI	98.4 %
without bROI	57.4 %

As Fig. 8 shows, partial extraction of the ROIs (Figs. 9(a)–(d)) and selection of epiphysis–metaphysis growth regions from the second or fourth fingers (Figs. 9(c)–(f)) are the two main reasons for failure when the bROIs are not employed. In our approach, because we first identify the enclosing areas in which the actual ROIs reside, the second case never occurs. In addition, instead of investigating the entire X-ray image by the Faster R-CNN, our approach allows only bROIs to be considered, thus more effectively capturing discriminative local features.

In our approach, the main cause of ROI extraction failure is due to the bROI extraction step. In particular, because the bROI extraction algorithm is based on conventional image processing techniques, its performance can be affected by image quality. For instance, it is highly likely that binarization cannot separate the hand area from the background of images with low contrast or intensities. To address the issue, we plan to develop a deep learning-based detection method for arbitrarily rotated bROIs.
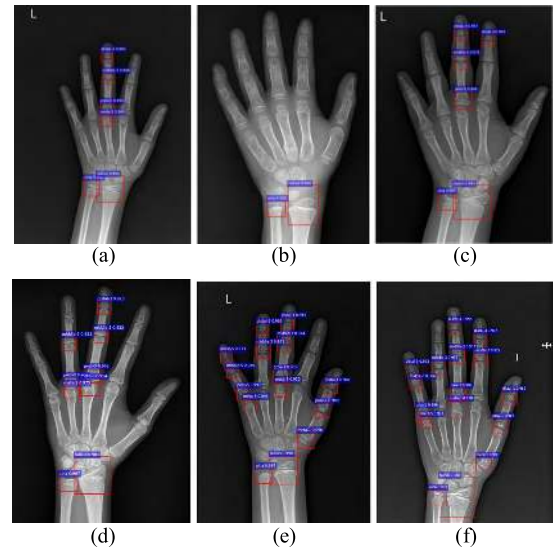
FIGURE 8. Failure cases of ROI extraction when the faster R-CNN without bROI is used. (a) 6 ROIs. (b) 2 ROIs. (c) 6 ROIs. (d) 9 ROIs. (e) 15 ROIs. (f) 16 ROIs.
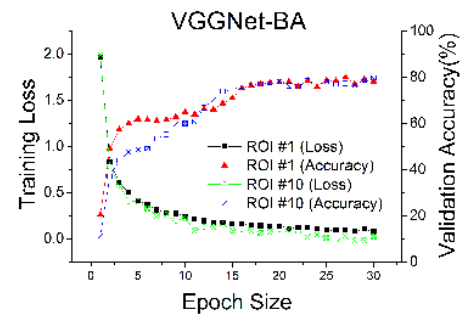
FIGURE 9. Loss and accuracy curves of the training process for example ROIs.

2) BA PREDCTION ACCURACY

Figure 9 shows the training loss and validation accuracy during the training of the VGGNet-BA models for two ROIs (i.e., the ulna and fifth metacarpal) as an example. In both cases, the curves converged properly with the hyperparameter settings, and it can therefore be concluded that the proposed models had learned as much as possible about the data. The training processes of CNN models for other ROIs also showed similar learning curve patterns.

To measure prediction accuracy, we used root mean squared error (RMSE) and the mean absolute error (MAE) as performance metrics. Their formulas are given below.

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^{n} |y_j - \hat{y}_j| \qquad \text{Eq.(2)}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^{n} (y_j - \hat{y}_j)^2} \qquad \text{Eq.(3)}$$

where n represents the number of test data items, and $\hat{y}_j$ and y_j denote the predicted BA and corresponding ground truth,

respectively. The MAE is a linear score which weights all individual differences equally in the average, whereas the RMSE gives a higher weight to large errors [36]. RMSE is most useful when large errors are particularly undesirable.

Table 6 shows the MAE and RMSE of the proposed deep-learning based BAA system according to age groups, and Table 7 shows the top-1 and top-2 prediction accuracies of the skeletal maturity levels of individual ROIs. In short, both the MAE and RMSE of the proposed system are relatively small and stable across all age groups and ROIs. For the ulna, all levels were included in training; its prediction accuracy is similar to those ROIs that only subsets of skeletal maturity levels are included for training. This result strongly suggests that, if a sufficient amount of training sample data is available, the inclusion of early stages of bone growth that were excluded in the experiments does not degrade the overall performance of the proposed BAA system.

TABLE 6. BA prediction accuracy per age group for the deep-learning based BAA system.

AGE (YEARS)	USING VGGNET-BA	
	MAE	RMSE
2–4	0.68	0.80
5–6	0.54	0.69
7–8	0.47	0.63
9–10	0.48	0.64
11–12	0.49	0.63
13–14	0.31	0.49
15–18	0.28	0.49
AVG.	0.46	0.62

TABLE 7. Top-1 and Top-2 accuracy of the deep-learning based BAA system.

ROI No.	USING VGGNET-BA	
	TOP-1	TOP-2
1	0.804	0.938
2	0.774	0.989
3	0.859	0.966
4	0.823	0.992
5	0.800	0.970
6	0.774	0.986
7	0.821	0.974
8	0.808	0.974
9	0.815	0.992
10	0.76	0.958
11	0.732	0.938
12	0.875	0.998
13	0.709	0.966
AVG.	0.796	0.972

Figure 10 shows the average Top-1 and Top-2 prediction accuracy for the skeletal maturity levels obtained using different loss functions. The VGGNet-BA with the class-weights-enabled loss function outperforms the one using conventional cross entropy loss function by 5.6% and 2.1% in Top-1 and Top-2 prediction accuracy, respectively. This result demonstrates that the application of the class-weights helps to reduce the negative effects of the data imbalance
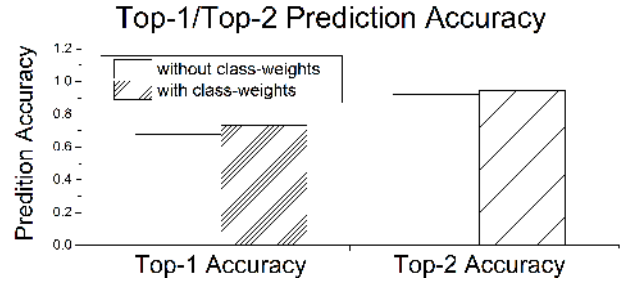

FIGURE 10. The comparative performances of the different loss functions.

to some extent. We also analyzed the confusion matrix of the VGGNet-BA to investigate its behaviors in more detail. As shown in Fig. 11, although all the confusion matrices of the components classifiers of the VGGNet-BA show strong diagonal patterns, implying that the labels predicted by VGGNet-BA are most often the correct skeletal maturity levels for all the ROIs, we find that there is still a performance difference between the levels owing to the imbalanced training samples. For instance, there were relatively few training samples at levels B, D, and H for the ulna, as shown in Table 2. The top-1 accuracies for levels B and D of the ulna were 32% and 17%, respectively and these values are considerably lower than the other levels that were trained with sufficient samples (see Fig. 11 (a)). However, VGGNet-BA achieved high prediction accuracies for level H. This is because there is a definite difference between level H and other levels in terms of the completion of the epiphyseal–diaphyseal fusion of the bone and, therefore, VGGNet-BA was able to learn the corresponding features, even with a small number of training samples for level H.

Although the dataset imbalance caused biased predictions, as shown in the experimental results, the reasons that the proposed BAA system using VGGNet-BA was able to perform acceptably are as follows. First, even in the case of predictions with very low levels of accuracy such as levels B and D, instead of using completely irrelevant levels, the proposed BAA system tries to evaluate the skeletal maturity as close to the ground truth as possible (see Table 6 and Fig. 10). Thus, errors in the RUS scores can be minimized. The second reason is attributed to the intrinsic strength of the TW3 method. That is, because the maturity level of each ROI is evaluated individually and independently, the TW3 method is robust to prediction errors among some bones.

We also measured prediction accuracies as well as inference times while changing the number of component classifiers in an ensemble (see Fig. 12). For this experiment, we used the same number of classifiers for all ROIs. The experimental results clearly show that the use of more component classifiers improved BA prediction accuracies, regardless of the classification network. However, above a certain number of classifiers, this improvement became marginal on both cases. The inference time is divided into three main components: the bROI extraction time (T_{broi}), actual ROI extraction time (T_{roi}), and classification time with
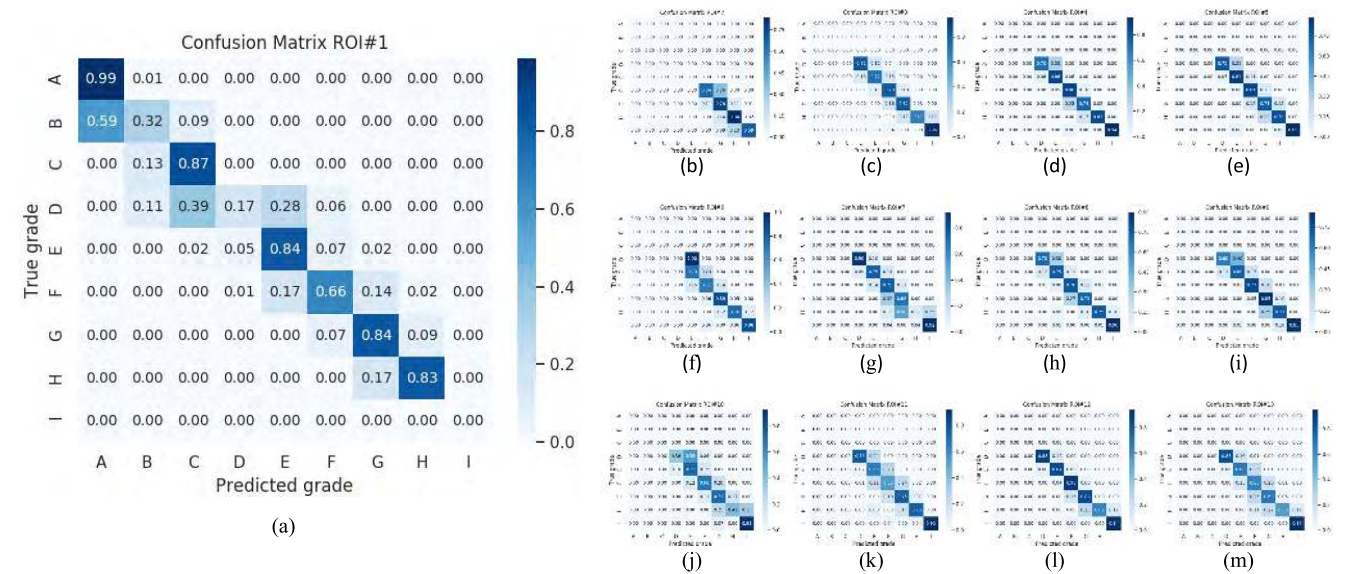
FIGURE 11. Normalized confusion matrix of VGGNet-BA. (a) ROI#1. (b) ROI#2. (c) ROI#3. (d) ROI#4. (e) ROI#5. (f) ROI#6. (g) ROI#7. (h) ROI#8. (i) ROI#9. (j) ROI#10. (k) ROI#11. (l) ROI#12. (m) ROI#13.
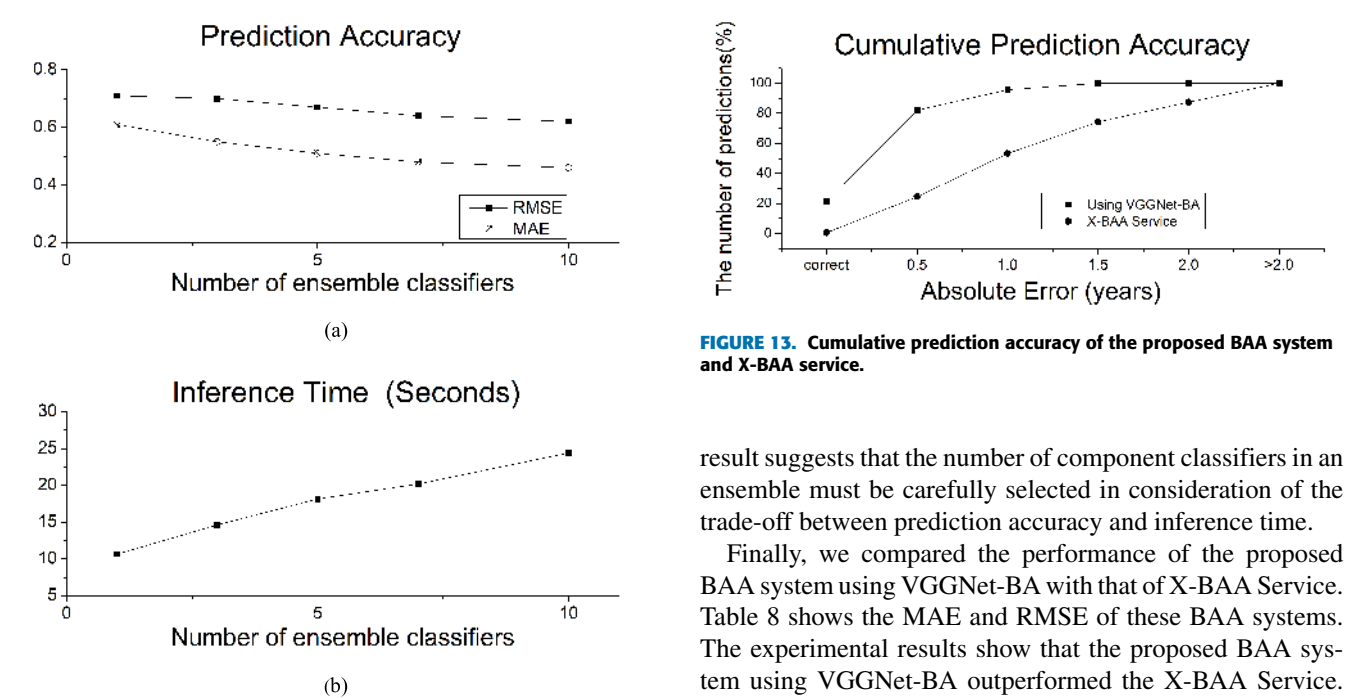

FIGURE 12. Sensitivity to the number of component classifiers in an ensemble. (a) Prediction accuracy for VGGNet_BA. (b) Inference time.

FIGURE 13. Cumulative prediction accuracy of the proposed BAA system and X-BAA service.

result suggests that the number of component classifiers in an ensemble must be carefully selected in consideration of the trade-off between prediction accuracy and inference time.

Finally, we compared the performance of the proposed BAA system using VGGNet-BA with that of X-BAA Service. Table 8 shows the MAE and RMSE of these BAA systems. The experimental results show that the proposed BAA system using VGGNet-BA outperformed the X-BAA Service.

N classifiers ($T_{classification}(N)$). Times T_{broi} and T_{roi} were 0.24 s and 8.63 s on average, respectively, and these times are not affected by the type of CNN network. However, $T_{classification}(1)$ was 1.54 s for VGGNet-BA. According to the figure, as the number of component classifiers increases, $T_{classification}$ increases linearly because the current prototype implementation does not allow GPU sharing among different CNN models so CNN models are run sequentially. This
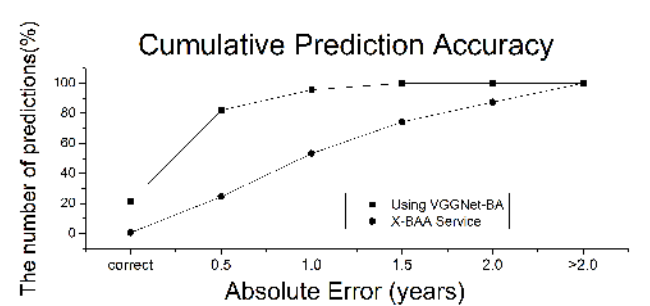
TABLE 8. Comparison of the proposed method with a commercial product.

AGE (YEARS)	USING VGGNET-BA		X-BAA SERVICE	
	MAE	RMSE	MAE	RMSE
2–4	0.61	0.74	1.92	1.97
5–6	0.53	0.63	1.30	1.93
7–8	0.43	0.56	1.12	0.85
9–10	0.45	0.58	1.11	1.34
11–12	0.49	0.61	0.61	0.75
13–14	0.32	0.36	1.58	1.76
AVG.	0.47	0.55	1.27	1.43

In particular, the magnitudes and variances of MAE and RMSE of the proposed BAA system are small and consistent, regardless of the dataset and age groups. This result suggests that the BA predicted by the proposed method is equal to the ground truth or the difference is marginal, as shown in Fig. 13. According to the figure, the accuracy (within one year of the ground truth) is 97.6% for the proposed system, whereas X-BAA Service achieved a comparable accuracy of 53.29%.

V. CONCLUSION

In this paper, we presented a TW3-based fully automated BAA system using deep neural networks. The core functions of the proposed system are the extraction of 13 ROIs and classification of the skeletal maturity levels of the ROIs. Additionally, we proposed techniques to address challenging problems in this research area. First, we introduced the concept of bROIs to reduce the area that Faster R-CNN must explore to extract an actual ROI. Second, we developed our own deep neural network that is customized for learning local discriminative features within small grayscale ROI images.

For the performance evaluation, we measured the MAE and RMSE of the proposed deep-learning based BAA system on a dataset of approximately 3,300 X-ray images. According to the results, the average top-1 and top-2 prediction accuracies for skeletal bone maturity levels for 13 ROIs were 79.6% and 97.2%, respectively. In addition, the MAE and RMSE for age prediction were 0.46 years and 0.62 years, respectively, and an accuracy within one year of the ground truth of 97.6% was achieved.

Future work lies in several areas. The proposed ROI extraction technique is hybrid in that bROI extraction is conventional image processing, whereas the extraction of actual ROIs uses deep learning techniques. Therefore, one direction of our future research is to develop deep learning-based algorithms and techniques for bROI extraction to accommodate diverse X-ray images with variations in intensity and contrast.

Owing to a lack of training samples, the current implementation did not use the early stages of bone growth in ROIs for deep learning, except for those of the ulna. Therefore, we intend to extend the current implementation by acquiring sufficient training samples for levels excluded in this study and training the system using all the levels of ROIs. In addition, we plan to improve the BA prediction accuracy to the level of professional radiology doctors. To this end, we will redesign component classifiers using state-of-the-art techniques, such as the inception module in GoogLeNet, residual learning in ResNet, and the dense block in DenseNet. These techniques provide innovations in the design of deep neural networks by enabling deeper networks with a smaller number of weights, thus leading to higher classification accuracies while avoiding the overfitting problem. Moreover, the configuration of the ensembles could be optimized. For instance, the ensembles could consist of completely different classifiers with different network topologies. Therefore, our future work may involve the exploration of various methods, algorithms, and techniques to improve the overall performance of our approach.

ACKNOWLEDGMENT

REFERENCES

[1] A. K. Poznanski, R. J. Hernandez, K. E. Guire, U. L. Bereza, and S. M. Garn, "Carpal length in children—A useful measurement in the diagnosis of rheumatoid arthritis and some congenital malformation syndromes," *Radiology*, vol. 129, no. 3, pp. 661–668, Dec. 1978.

[2] V. Gilsanz and O. Ratib, *Hand Bone Age: A Digital Atlas of Skeletal Maturity*, 1st ed. New York, NY, USA: Springer, 2005.

[3] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in X-Ray images," *Med. Image Anal.*, vol. 36, pp. 41–51, 2017.

[4] W. W. Greulich and S. I. Pyle, "Radiographic atlas of skeletal development of the hand and wrist," *Amer. J. Med. Sci.*, vol. 238, no. 3, p. 393, 1959.

[5] H. Goldstein, J. M. Tanner, M. Healy, and N. Cameron, *Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method)*, 3rd ed. New York, NY, USA: Harcour, 2001.

[6] K. Khan and A. S. Elayappen, "Bone growth estimation using radiology (Greulich–Pyle and Tanner–Whitehouse methods)," in *Handbook of Growth and Growth Monitoring in Health and Disease*, V. R. Preedy, Ed. New York, NY, USA: Springer, 2012, pp. 2937–2953.

[7] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[8] M. Ghafoorian *et al.*, "Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities," *Sci. Rep.*, vol. 5, Jul. 2017, Art. no. 5110.

[9] J.-P. Charbonnier, E. M. van Rikxoort, A. A. A. Setio, C. M. Schaefer-Prokop, B. van Ginneken, and F. Ciompi, "Improving airway segmentation in computed tomography using leak detection with convolutional networks," *Med. Image Anal.*, vol. 36, 52–60, Feb. 2017.

[10] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.

[11] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[12] H. Lee *et al.*, "Fully automated deep learning system for bone age assessment," *J. Digit. Imag.*, vol. 30, no. 4, pp. 427–441, 2017.

[13] M. Chen, "Automated bone age classification with deep neural networks," Stanford Univ., Stanford, CA, USA, Tech. Rep, 2016. [Online]. Available: http://cs231n.stanford.edu/reports/2016/pdfs/310_Report.pdf

[14] R. Bakthula and S. Agarwal, "Automated human bone age assessment using image processing methods—Survey," *Int. J. Comput. Appl.*, vol. 104, no. 13, pp. 1–10, 2014.

[15] M. Mansourvar, M. A. Ismail, T. Herawan, R. G. Raj, S. A. Kareem, and F. H. Nasaruddin, "Automated bone age assessment: motivation, taxonomies, and challenges," *Comput. Math. Methods Med.*, vol. 2013, Oct. 2013, Art. no. 391626.

[16] H. H. Thodberg, S. Kreiborg, A. Juul, and K. D. Pedersen, "The BoneXpert method for automated determination of skeletal maturity," *IEEE Trans. Med. Imag.*, vol. 28, no. 1, pp. 52–66, Jan. 2009.

[17] J. Kim *et al.*, "Computerized bone age estimation using deep learning based program: Evaluation of the accuracy and efficiency," *Amer. J. Roentgenology*, vol. 209, no. 6, pp. 1374–1380, 2017.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CPVR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[19] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[20] K. Simonyan and A. Zisserman. (Sep. 2017). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: https://arxiv.org/abs/1409.1556

[21] J. Seok, B. Hyunm J. Kasa-Vubu, and A. Girard, "Automated classification system for bone age X-ray images," in *Proc. IEEE Int. Conf. Syst. Man, Cybern.*, Seoul, South Korea, Oct. 2012, pp. 208–213.

[22] A. Tristán-Vega and J. I. Arribas, "A radius and ulna TW3 bone age assessment system," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 5, pp. 1463–1476, May 2008.

[23] B. Fischer, P. Welter, R. W. Günther, and T. M. Deserno, "Web-based bone age assessment by content-based image retrieval for case-based reasoning," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 7, no. 3, pp. 389–399, 2012.

[24] M. Harmsen, B. Fischer, H. Schramm, T. Seidl, and T. M. Deserno, "Support vector machine classification based on correlation prototypes applied to bone age assessment," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 1, pp. 190–197, Jan. 2013.

[25] M. Kashif, S. Jonas, D. Haak, and T. M. Deserno, "Bone age assessment meets SIFT," *Proc. SPIE*, vol. 9414, Mar. 2015, Art. no. 941439.

[26] K. Hanbin, "Automated bone age assessment for ulna and radius using TW3," M.S. thesis, Dept. Comput. Sci., Kyonggi Univ., Kyonggido, South Korea, 2016.

[27] M. Alcina, A. Lucea, M. Salicrú, and D. Turbón, "Reliability of the Greulich and Pyle method for chronological age estimation and age majority prediction in a Spanish sample," *Int. J. Legal Med.*, vol. 132, no. 4, pp. 1139–1149, Jul. 2015.

[28] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process.*, 2015, pp. 91–99.

[29] OpenCV Library. (Jan. 2018). *Image Processing*. [Online]. Available: http://opencv.org

[30] (Oct. 2018). *Faster R-CNN (Python implementation)*. [Online]. Available: https://github.com/rbgirshick/py-faster-rcnn

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1026–1034.

[32] R. Bonab and F. Can. (2017). "Less is more: A comprehensive framework for the number of components of ensemble classifiers." [Online]. Available: https://arxiv.org/abs/1709.02925

[33] Y. J. Oh, "Development and evaluation of semi-automatic bone age estimation method," M.S. thesis, Ewah Womans Univ., Seoul, South Korea, 2011.

[34] (Jan. 2019). *Learning From Imbalanced Data*. [Online]. Available: https://www.jeremyjordan.me/imbalanced-data/

[35] Scikit-learn. (Feb. 2019). *Class Weights*. [Online]. Available: https://scikit-learn.org

[36] (Jan. 2018). *MAE and RMSE—Which Metric is Better?* [Online]. Available: https://medium.com/human-in-a-machine-world/mae-and-rmse-which-metric-is-better-e60ac3bde13d/

[37] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4700–4708.

SUNG JOON SON was born in Seoul, South Korea, in 1989. He received the B.S. degree from the School of Biomedical Engineering, Korea University, Seoul, in 2015. He is currently pursuing the Ph.D. degree with the Graduate School of Convergence Science and Technology, Seoul National University. His current research interests include computer vision and deep learning.

YOUNGMIN SONG received the B.S. and M.S. degrees in computer science from Kyonngi University, Suwon, South Korea, in 2016 and 2018, respectively, where he is currently a Research Staff Member of the Department of Computer Science. His research interests include deep learning, image processing, and medical image analysis.

NAMGI KIM received the B.S. degree in computer science from Sogang University, South Korea, in 1997, and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, in 2000 and 2005, respectively. From 2005 to 2007, he was a Research Staff with Samsung Electronics. Since 2007, he has been a Faculty Member with Kyonggi University, South Korea, where he is currently an Associate Professor with the Department of Computer Science. His research interests include sensor systems, wireless systems, and mobile communication.

YOUNGHAE DO received the M.S. degree from Rutgers University, in 1999, and the Ph.D. degree from Arizona State University, in 2004. Since 2007, he has been a Professor with the Department of Mathematics, Kyungpook National University, South Korea. He was the Director of Research Center called Computation and Methodology in Applied Fluid Dynamics, from 2009 to 2013, which is one of the world-class university projects. His research interests include nonlinear dynamics, chaos, game theory, network theory, and bio-medical mathematics. Since 2015, he has been serving as an Editorial Member of *Scientific Reports*.
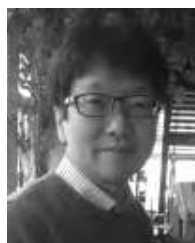
NOJUN KWAK was born in Seoul, South Korea, in 1974. He received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering and Computer Science, Seoul National University, Seoul, in 1997, 1999, and 2003, respectively. From 2003 to 2006, he was with Samsung Electronics. In 2006, he joined Seoul National University, as a BK21 Assistant Professor. From 2007 to 2013, he was a Faculty Member with the Department of Electrical and Computer Engineering, Ajou University, Suwon, South Korea. Since 2013, he has been with the Graduate School of Convergence Science and Technology, Seoul National University, where he is currently a Professor. His current research interests include feature learning by deep neural networks and their applications in various areas of pattern recognition, computer vision, and image processing.

MU SOOK LEE received the M.D. degree from Keimyung University, South Korea, in 2002, and the M.S. degree from Yonsei University, South Korea, in 2008. She is currently a Radiologist with the Human Medical Imaging Center, Seoul, South Korea. In 2002, she passed a licensing examination in radiology in Korea. Her research interests include pediatric radiology and interventional radiology.

BYOUNG-DAI LEE received the B.S. and M.S. degrees in computer science from Yonsei University, South Korea, in 1996 and 1998, respectively, and the Ph.D. degree in computer science and engineering from the University of Minnesota, Minneapolis, USA, in 2003. Before joining Kyonngi University, South Korea, he was a Senior Engineer with Samsung Electronics, Co., Ltd., from 2003 to 2010. He is currently an Associate Professor with the Department of Computer Science, Kyonngi University. During the period, he has participated in many commercialization projects related to mobile broadcast systems. In 2014, he founded a venture company, BLEE, Co., where he is currently the CEO. His research interests include machine learning, deep learning, and medical image analysis.

• • •