

RESEARCH ARTICLE

Deep Learning and Computer Vision Techniques for Automated Total Hip Arthroplasty Planning on 2-D Radiographs

MINWOO KIM¹, IL-SEOK OH^{1,2}, AND SUN-JUNG YOON^{3,4}¹Division of Computer Science and Engineering, Jeonbuk National University, Jeonju 54896, South Korea²Center for Advanced Image Information Technology, Jeonbuk National University, Jeonju 54896, South Korea³Department of Orthopedic Surgery, Medical School, Jeonbuk National University, Jeonju 54907, South Korea⁴Research Institute of Clinical Medicine of Jeonbuk National University–Biomedical Research Institute of Jeonbuk National University Hospital, Jeonju 54907, South Korea

Corresponding author: Sun-Jung Yoon (sunjungyoon@jbnu.ac.kr)

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education under Grant 2021R1I1A3060112; in part by the Grant of the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea, under Grant HR22C1832; and in part by the Fund of Biomedical Research Institute, Jeonbuk National University Hospital.


ABSTRACT Preoperative planning is mandatory for successful total hip arthroplasty (THA). In planning, the operating surgeon should decide the best type and size of THA components for the patient. However, most digital templating software only simulates acetate templating by overlaying the shape of the prosthesis components on a radiograph; the selection and positioning of the prostheses are performed manually depending on the operator's experience. Determining the optimal type and size of THA components is a repetitive and time-consuming task for digital and acetate templating. This study proposes a novel approach to automatically select and position THA components that are most suitable for the patient's bone anatomy. The approach consists of two phases: segmenting a hip anteroposterior (AP) radiographic image into five predefined anatomical regions using a fully convolutional neural network, and estimating the optimal sizes and positions of THA components using deep learning and computer vision technology. The experiments demonstrated that the accuracy of acetabular and femoral component size prediction within one size error was 78.9% and 70.9%, respectively. Compared with meta-analysis results from previous studies, our results are close to human level. An automated digital templating prototype system was developed using our research results and tested in a clinical setting to evaluate field adaptability. These processes are introduced in this study.

INDEX TERMS Artificial intelligence (AI), automated 2-D templating, hip joint segmentation, implant size estimation, prostheses recommendation, total hip arthroplasty (THA).

I. INTRODUCTION

An arthroplasty that replaces the joint with a prosthesis is required when a patient has end-stage arthritis of the hip joint. Preoperative planning is of paramount importance to obtain reproducible results in modern total hip arthroplasty (THA) [1], [2]. In THA planning, an orthopedic surgeon decides the

shape and size of the prostheses (AC; acetabular component, FC; femoral component) most suitable to the patient's bone anatomy. The preoperative planning process usually involves acetate or digital templating on medical images such as 2-D and 3-D radiography, and computed tomography (CT) [2], [4], [5]. Digital templating enables an operator to manually select from different types and sizes of prostheses and validate them by overlaying them on images, the acetate templating process performed in a digital environment. That is,

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

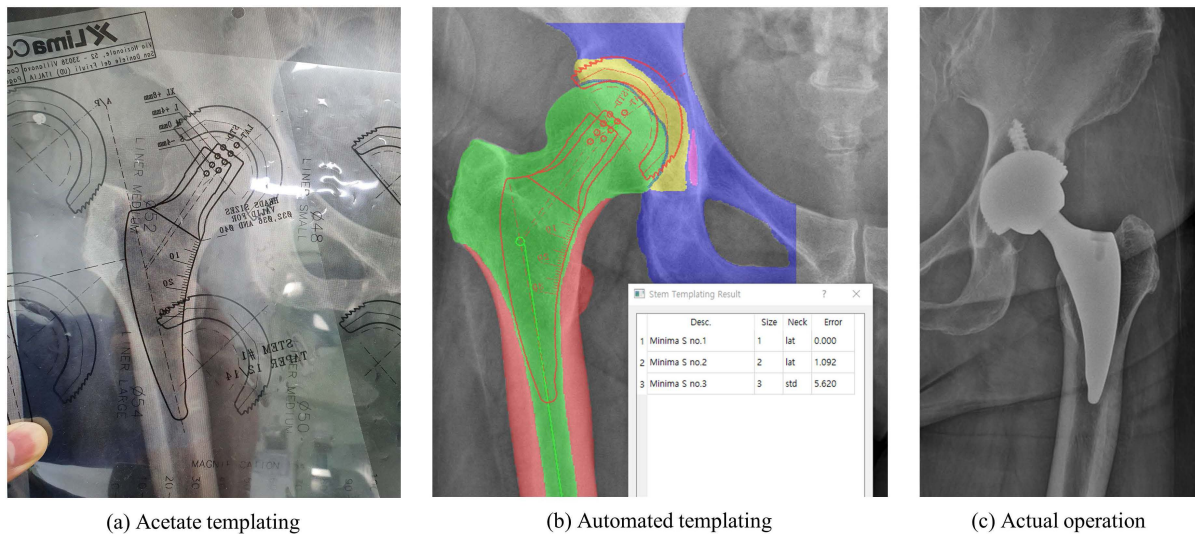


FIGURE 1. Comparison of conventional acetate templating and proposed automated templating. (a) and (b) are samples of conventional acetate templating and proposed AI-based automated templating, respectively, on the side opposite the side to be operated on. (c) is the actual operation result on the same patient.

the selection and positioning of the prostheses are performed manually depending on the operator's experience.

With recent breakthroughs in computer vision and deep learning technologies, automatic decision-making for many medical imaging problems can be achieved with practical performance [6], [7], [8], [9], [10], [11]. Segmentation problems are more complicated than classification problems [12]. However, the advent of deep learning has changed the overall picture, making high-performance segmentation easier [13], [14], [15], [16]. The use of deep learning in natural image segmentation has led to significant advances in medical image segmentation [17], [18], [19], [20]. In 2015, U-net was proposed for segmenting medical images, becoming the most popular tool for medical image segmentation [8], [10], [21], [22].

Arthroplasty can benefit from these efforts and progress in medical imaging, leading to fully automatic decision-making for prosthesis selection. However, our literature survey found few studies related to decision automation [50], [54], [60]. Most of the published studies have investigated hip joint segmentation or compared the accuracy and reliability of manual preoperative THA planning [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59]. A detailed literature survey is presented in Section 2.

Fig. 1(a) illustrates a conventional 2-D acetate templating procedure in which a doctor manually chooses the prostheses to fit the patient. First, the displayed magnification of the hip anteroposterior (AP) radiographic image was adjusted to match the magnification of the acetate template. By overlaying the template on the hip AP image, the optimal shape, size, and position of the components are estimated by the surgeon [1], [2], [3]. Because the morphology of the human acetabulum and proximal femur is diverse, manually

selecting the most suitable prosthesis for the hip joint is inefficient and time-consuming. Moreover, owing to differences in magnification, applying the template provided by the manufacturer to the digital image displayed on a computer monitor was inconvenient, and simultaneously comparing multiple templates from different families was impossible. If the process is automated and augmented by deep learning with computer vision technology, the operating surgeon can obtain higher-level surgical planning results using artificial intelligence (AI)-based automated preoperative templating to select the proper prosthesis, restore natural biomechanics, and decrease operation time, ultimately leading to greater patient satisfaction.

This study proposes a novel approach for intelligent decision-making on patient-specific prosthesis selection in hip AP radiographs, as shown in Fig. 1(b). The THA components most suitable for the patient's bone anatomy are automatically selected and located, and a list is presented to the operating surgeon in order of suitability. We adopted a conventional thinking process by an experienced arthroplasty surgeon. Our approach has two stages: segmenting the relevant regions from an input radiographic image and fitting prostheses on the segmented regions, as shown in Fig. 2. We annotated five regions in a pixel-wise manner for 151 hip AP images for segmentation. Fig. 3 shows an example of the annotation of the five regions. Then, we trained U-net with the annotated data to segment regions. The second phase fitted prostheses. To predict the size of the AC, we employed transfer learning, which reused the feature extraction path (first half) of the U-net learned for segmentation. We adopted various computer vision algorithms and developed several rules for selecting the FC size and placement of the AC and FC in the segmented regions. Deep learning trained-templating software was built for a field test incorporating the modules described in this study. The accuracy of the proposed

automated templating method was retrospectively evaluated in 128 patients who underwent THA. In addition, the system is virtually located in an actual operating theater at Jeonbuk National University Hospital. The clinical feasibility and applicability are being evaluated.

In summary, the main contribution of this work is three-fold. (1) We propose an AI-based automated preoperative THA planning system. To the best of our knowledge, this is the first attempt at disclosing an entire workflow that automatically determines the size and location of ACs and FCs. (2) To increase applicability and accessibility, we use simple radiographic images. Previous studies have focused more on 3-D CT and MR data than on 2-D radiographic images. Simple radiographs have several advantages over MRI and CT images including real-time insight, cost-effectiveness, and intraoperative reproducibility. (3) The proposed automated planning system is applicable to all prostheses manufacturers. That is, various types of acetate and digital templates can be imported and reviewed simultaneously, regardless of the magnification of each template.

II. RELATED RESEARCH

Accurate hip joint segmentation from medical images (radiographs, CT, MR) is a prerequisite for computer-assisted diagnosis and preoperative planning in orthopedic surgeries [23]. In this regard, automatic segmentation of the pelvis and femur from 3-D data has been a major concern in recent decades. These methods can be roughly categorized as supervised or unsupervised approaches based on prior knowledge. Unsupervised approaches that do not use prior knowledge can be further categorized as region growing [24], thresholding [25], [26], graph-cut [27], surface fitting [28], and harmonic field approaches [29]. Supervised approaches using prior knowledge can be further categorized as statistical shape models (SSM) [30], [31], [32], active shape models (ASM) [33], atlas-based [34], [35], and patch-based approaches [36]. In recent years, deep learning-based approaches have also been used [37], [38], [39]. Deniz *et al.* devised a deep learning-based method that exploited the powerful feature extraction ability of CNNs and the advantage of end-to-end learning ability [38]. They compared two types of U-nets: 2-D U-nets applied to each 2-D slice of the MR image, combining the results, and 3-D U-nets applied once to the MR image. Zeng *et al.* proposed a deeply supervised 3-D U-net-like fully convolutional network for femur segmentation from MR images [37]. They introduced multilevel deep supervision and partial transfer learning to boost training efficiency. Chen *et al.* presented a 3-D feature-enhanced network to achieve fast and accurate femur segmentation from CT images [39]. The feature-enhanced network used skip connections to fuse multiscale feature maps as a U-net and fused the edge detection task to optimize femur segmentation through joint learning.

Automatic hip joint segmentation of 2-D radiographs has been studied less than segmentation of 3-D images. Conventional methods find contours of the pelvis and femur using

SSM [40], [41], [43], ASM [44], shortest path [45], and generic 3-D model projection and registration [46]. Traditionally, segmentation of 2-D radiographic images has been beneficial but rarely studied owing to its difficulties, including poor and non-uniform image contrast, noise, occlusions, and overlap of neighboring structures. Recently, deep learning-based methods widely used for medical image segmentation have been applied to hip joint segmentation, with excellent results [47], [48]. Shen *et al.* proposed a pure dilated residual U-net that improved the accuracy and convergence of vanilla U-net [21] in segmenting femurs and tibias [47]. Rouzrokh *et al.* proposed a method for the automated radiographic measurement of AC inclination and version after THA [48]. They segmented the AC and ischial tuberosities using U-net. The AC angle was measured on the predicted segmentation map using image processing techniques.

Most published studies directly related to preoperative templating of THA have evaluated the accuracy and reliability of 2-D acetate and 2-D/3-D digital templating by comparing preoperative templating results with actual surgical results, prospectively [49], [50], [51] and retrospectively [52], [53], [54], [55], [56], [57]. Systematic reviews and meta-analyses of these studies have also been published [4], [58], [59]. In summary, 1) 2-D and 3-D digital templating are reliable techniques for preoperative THA planning; 2) templating results are accurate in order of 3-D digital, 2-D digital, and 2-D acetate; and 3) templating is less accurate with uncemented prostheses than with cemented prostheses. For preoperative THA planning, 2-D digital templating is still a standard despite the superior accuracy of 3-D templating, as it involves a smaller radiation dose and lower imaging and software costs [59].

Few studies have been published on automated preoperative THA templating. In an early study, Otomaru *et al.* proposed automated AC planning using 3-D CT data [60]. They used two statistical atlases, a combined pelvis and cup statistical shape model (PC-SSM), and a statistical map of bone thickness (SM-BT) to detect the pelvis surface from CT data. The AC size and position were estimated from the segmentation results. Huo *et al.* and Ding *et al.* presented prospective and retrospective studies on the value of AI-based 3-D preoperative planning software (AI-HIP) for THA [50], [54]. AI-HIP consists of a segmentation module based on U-net for segmentation of the pelvis and femur from a 3-D CT, a landmark detection module based on a neural network, and an automatic search engine based on big data and reinforcement learning for proper prosthesis selection. They reported that AI-HIP with 3-D CT data is more accurate and reliable for preoperative THA planning than 2-D digital templating and 3-D mimics software.

III. METHODS

This section describes the deep learning architecture, learning schemes, and rule-based computer vision techniques for segmenting osseous structures and fitting ACs and FCs, as shown in Fig. 2. The datasets are also described as the data and

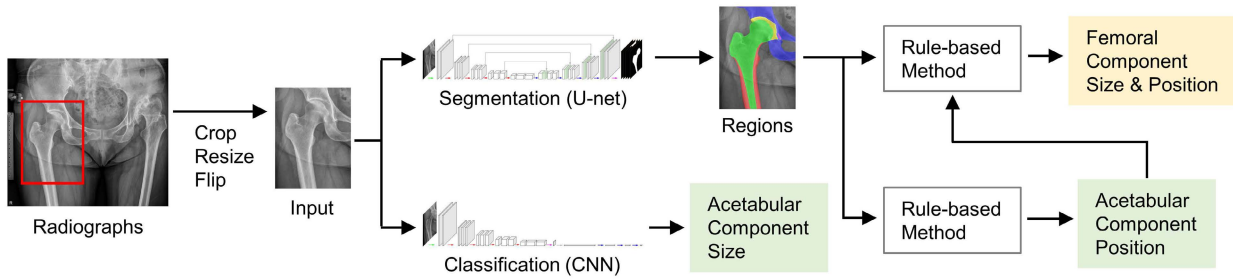


FIGURE 2. Overview of proposed AI-based automated preoperative templating procedure. We trained U-net to segment five predefined anatomical regions on hip AP images. To predict the size of the acetabular component (AC), we employ transfer learning, which reuses the feature extraction path (first half) of the U-net learned for segmentation. We use various computer vision algorithms and developed several rules for selecting the femoral component (FC) size and placement of the AC and FC in the segmented regions.

augmentation methods are significant in designing and implementing deep learning models.

A. DATASETS

To prepare a training set for a deep learning model to achieve osseous structure segmentation, we manually segmented the contralateral side of a hip AP radiographic image into five regions, as shown in Fig. 3, using Adobe Photoshop. Unilateral hip and pelvis AP images were cropped into the proximal femur and hemipelvis. The proximal femur was divided into two regions: HC (head and canal; femoral head, medullary canal, and greater trochanter) and C (cortex; inferior neck cortex, lesser trochanter, and diaphyseal cortex). The hemipelvis was divided into three regions: pelvis (P), joint space (J), and teardrop (T). Each region was stored in a separate map; thus, a hip radiograph had five maps as the ground truth. A total of 151 AP images of THA patients were labeled, 115 for training and 36 for validation.

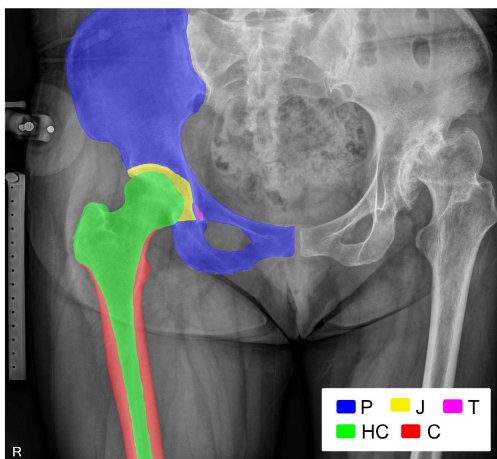


FIGURE 3. Five regions of radiographic image: P; Pelvis, T; Teardrop, J; Joint space and saucily, HC; Femoral head and canal, C; Cortex.

To alleviate overfitting caused by data scarcity, we augmented the image 100 times to obtain 15100 images for the segmentation task. We placed the cropping box, including the main components of the proximal femur and hemipelvis

(red box in Fig. 2), and randomly translated, rotated (-7° – 7°), and scaled (0.85–1.15) the cropping box. The box was cropped and resized as input for the network. If the osseous structures to be segmented were on the right side (left hip and pelvis) of the image, the cropped image was flipped such that the model inputs were consistent.

To predict the AC size, we chose 110 images from 151 images labeled with the AC size information used in the actual operation, and placed a ruler inside the image. We calculated the actual dots per inch (dpi) to determine the magnification using the ruler. Each radiographic image was labeled with the AC sizes using medical records collected retrospectively. The distribution of the labeled AC sizes is shown in Table 1. The number of images was highly unbalanced between classes. Of the ten AC sizes, two sizes had no images, and sizes 44, 56, and 58 had only three, two, and one image(s), respectively.

TABLE 1. Distribution of acetabular component sizes in dataset used for size prediction.

Class (Component Size)	Number of Images	Training		Validation	
		Original	Augmentation	Original	Augmentation
44	3	2	930	1	10
46	13	10	930	3	30
48	19	16	930	3	30
50	29	26	930	3	30
52	12	10	930	2	20
54	31	28	930	3	30
56	2	1	930	1	10
58	1	0	930	1	10
60	0	0	930	0	0
62	0	0	930	0	0
Total	110	93	9300	17	170

To address the imbalanced data, we generated synthetic samples of nine other sizes from an image with size i , as shown in Fig. 4. That is, ten samples (nine synthetic samples and one real sample) are generated from one image for ten sizes. Our AC size estimation model assumes that the input image patches have a fixed resolution (70 dpi) to predict their actual size in millimeters. Thus, we generated synthetic samples for the other nine sizes by changing the resolution

of the input image patches. For the synthetic sample with target size j from the image with AC size i , the input resolution was calculated as $70 * j/i$ dpi. Patch cropping was performed by slightly shifting and rotating (-7° – 7°) to further augment the images. The amount of shifting and rotation was randomly chosen within a small range. Random cropping was performed ten times. A total of 100 patches belonging to ten classes were generated from an image. The cropped patches were resized and flipped for input into the network. Fig. 4 shows an example with an original resolution of 200 dpi and an actual AC size of 46.

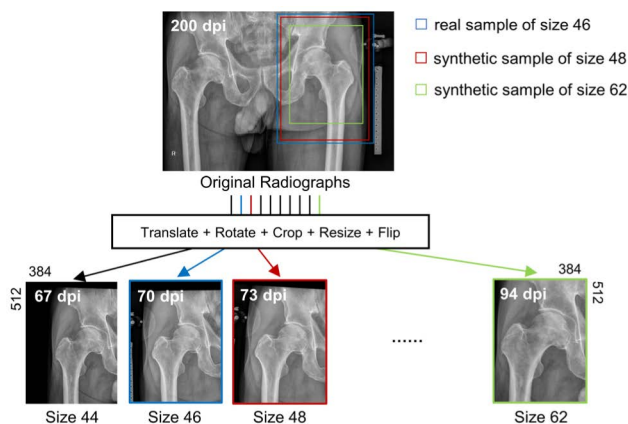


FIGURE 4. Data augmentation for acetabular component size estimation. A sample of a real component size and nine synthetic samples of different sizes are produced from one image.

To evaluate our method for predicting AC and FC size, we prepared a test set with the ground truth from the surgeon during the actual operation. Hip AP radiographs were collected before surgery from 128 THA patients; AC and FC sizes were extracted from the medical records. The collection was not used, and different patients were used for training and validating our deep learning models. Sixty-five patients had end-stage arthrosis of the left hip; the remaining 63 patients had end-stage arthrosis of the right hip.

B. SEGMENTATION OF HEMIPELVIS AND PROXIMAL FEMUR

We used U-net for segmentation of a hip AP radiographic image into five regions, as shown in Fig. 5. U-net receives a medical image as input and outputs a segmentation map. The input to our network was a patch cropped and resized from the radiographic image shown in the red box in Fig. 2. The resized input patch was 512×384 in size. The output of the network was five feature maps 512×384 in size; the output was represented as a $5 \times 512 \times 384$ tensor, with each tensor representing the C, HC, P, J, and T regions. Cropping is performed by the user in the test phase, and U-net performs segmentation.

The architecture of U-net is shown in Fig. 5; the contracting and expanding paths produce symmetry. The role of the contracting path is to extract rich features from the input image.

The feature map was maintained at a constant size through padded convolution at each level. A total of 1024 feature maps were generated, each with a size of 32×24 at the bottom level. The expanding path uses the generated feature maps to synthesize segmentation maps, represented as a $5 \times 512 \times 384$ tensor. It has skip connections to transfer the high-resolution features of the contracting path to the expanding process. The transferred feature maps are concatenated to the maps generated by the layers on the expanding path. Skip connections are the core concept of U-net, and are critical in generating high-quality segmentation maps for medical images. For U-net training, binary cross-entropy was used as the loss function; a stochastic gradient descent (SGD) optimizer was used with a learning rate of 0.0001, momentum of 0.99, and batch size of 1.

C. ACETABULAR COMPONENT SIZE PREDICTION AND POSITIONING

Predicting the AC involves two steps: estimating the optimal size using a convolutional neural network (CNN), and using rule-based landmark detection on the segmentation results for proper placement. Component size estimation is performed as an ordinal classification problem. We used transfer learning, reusing the parameter value learned from training the segmentation U-net to train the CNN for AC size estimation. We considered the contracting path of the learned U-net shown in Fig. 5 and attached the fully connected layers for classification. Fig. 6 illustrates the architecture of CNN classification. The output layer had ten nodes corresponding to ten AC sizes (44, 46, 48, . . . , 62), and used softmax as the activation function. For training, CNN cross-entropy was used as the loss function, and the SGD optimizer was used with a learning rate of 0.0001, momentum of 0.9, and batch size of 10.

After determining the AC size, the position and placement were calculated based on the segmentation information. To this end, two landmarks (L1 and L2) were automatically determined using segmented regions, shown as red points in Fig. 7(b). L1 was the most lateral point of the J (joint space) region. L2 was the lowest point of the J region, adjacent to the HC region (femoral head and canal). The baseline of the AC was aligned with the line connecting L1 and L2 such that their centers coincided.

D. FEMORAL COMPONENT SIZE PREDICTION AND POSITIONING

Determining the FC size and placement was based on the segmentation results. The algorithm was designed using three rules relevant to the actual operation.

- 1) The central axis of the FC should be aligned with that of the proximal femur.
- 2) The FC should not intrude into the C (cortex) region extracted by segmentation and should be placed deep in the cancellous HC region.
- 3) The FC trunk should be sufficiently close to the cup center.

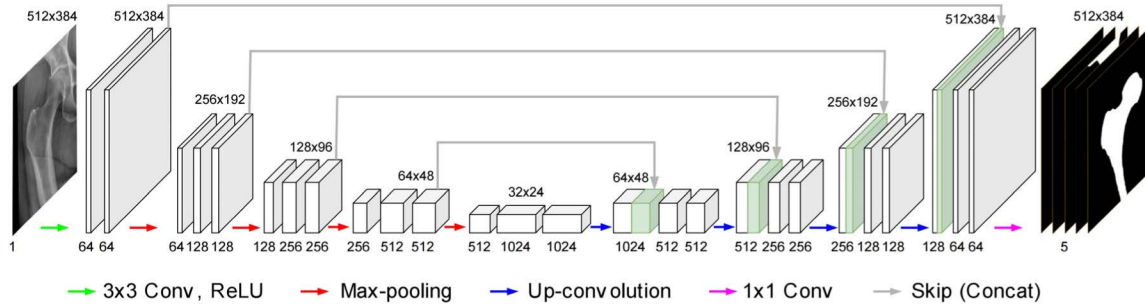


FIGURE 5. U-net for segmenting hip AP radiograph image into five regions.

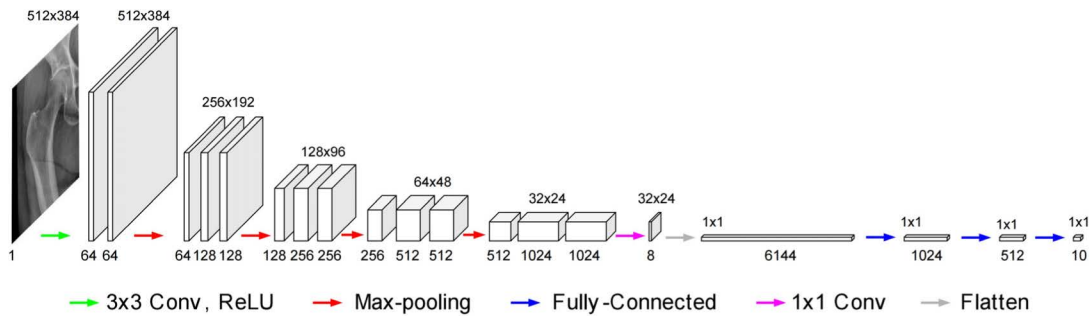


FIGURE 6. CNN for classification of acetabular component size prediction.

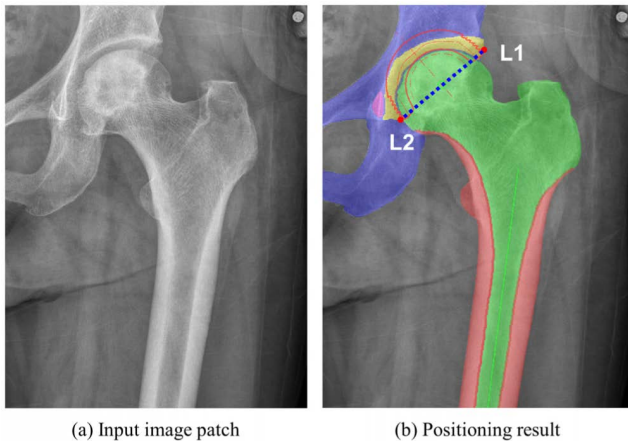


FIGURE 7. Position and placement calculation for acetabular component. Two landmarks (L1 and L2) were automatically determined using segmented regions and serve as the basis for placing the acetabular component.

To satisfy the first rule, we determined the central axis of the proximal femur. The C region is the diaphyseal cortex of the proximal femur; its central axis is regarded as the central axis of the femur. Applying the distance transform operation to C [61], we obtained the distance map in Fig. 8(a). The distance transform computes the shortest distance from each pixel in region C. We obtained the green area in Fig. 8(a) by extracting the ridge points from the distance map. We obtained the central axis of the femur by fitting a straight line to the ridge points, indicated by the

blue line in Fig. 8(b). The second rule was used to insert the FC deep inside the cancellous bone region of the medullary canal (HC). By aligning the central axes of the FC and femur, we slid the FC as deep as possible without intruding into the diaphyseal cortex (C) region. Fig. 8(b) shows the completed insertion.

The third rule was used to evaluate the placement of the FC in terms of the matching error. The matching error was measured by calculating the angle between two straight lines, $O_{stem}E_{neck}$ and $O_{stem}O_{cup}$. Fig. 8(c) illustrates the matching error calculation process. The algorithm exhaustively searches for all available FCs of different sizes and shapes. The FCs are recommended to the operating surgeon in order of smallest matching error.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENT ENVIRONMENT

The experiments were conducted on an AMD Ryzen7 2700X with a 3.7 GHz CPU, 32 GB of memory, and a GTX 1080 Ti GPU. We aimed to quantitatively and qualitatively evaluate the quality of bone segmentation and AC and FC size prediction. The deep learning component was implemented using Keras with a TensorFlow backbone.

Hip AP radiography was taken at a distance of approximately 120 cm from the X-ray tube to the cassette, and the osseous structure was magnified by approximately 110%. To determine the correct magnification of the radiographs, a ruler was placed at the center level of the femur. To offset the femoral version, preoperative templating was filmed in the

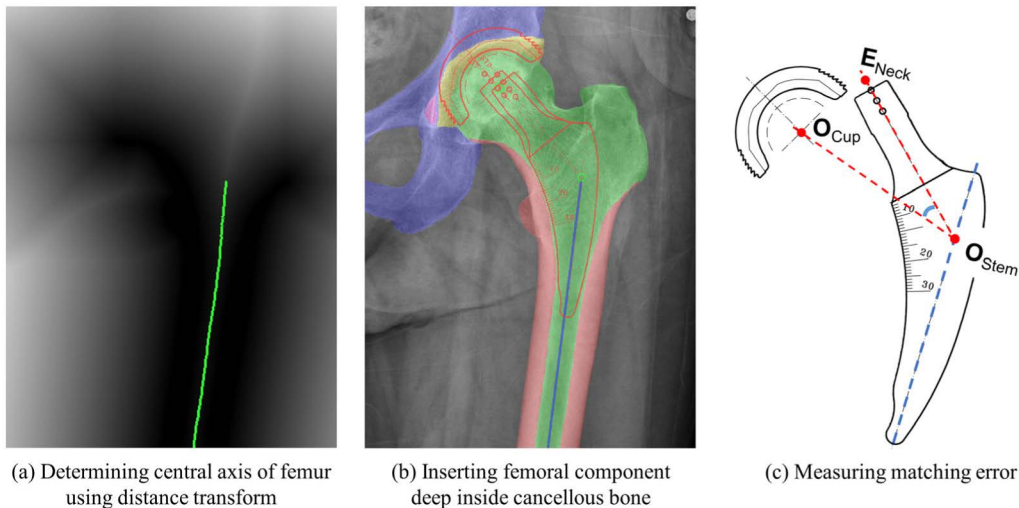


FIGURE 8. Determining and placing femoral component of prosthesis. The femoral component is inserted along the central axis of the femur and evaluated for how close the trunk is to the center of the acetabular component.

state of internal rotation maximally with the lower extremity. In addition, internal rotation is difficult with a diseased hip; thus, templating was performed on the side opposite to the side to be operated on.

The prostheses used for performance evaluation was the AC (DELTA PF, LimaCorporate, Italy) and FC (Minima S and MASTER SL, LimaCorporate, Italy), as shown in Fig. 9. The AC has 12 sizes from 44 to 66 in 2 unit increments, but a size of 64 or higher is rarely used; therefore, we used 10 steps from 44 to 62 for templating. The FC Minima S and MASTER SL stems consist of 12 and 13 steps, respectively, and are all considered in templating. All prostheses used in this study were uncemented.

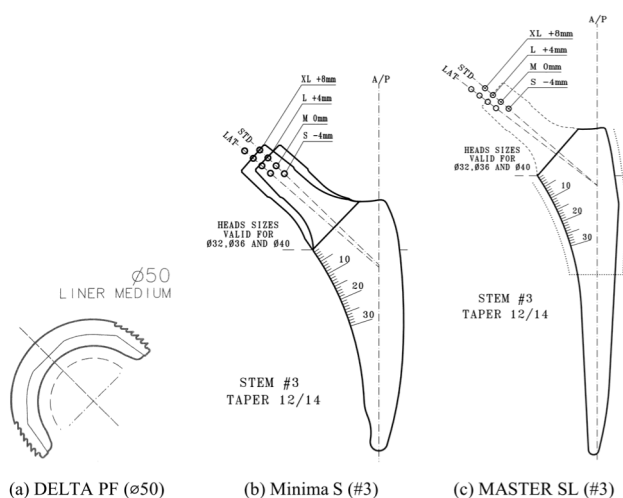


FIGURE 9. Design samples of acetabular and femoral components used for evaluation. The template from a medical device company is a simplified outline of the prosthesis; the length and offsets are indicated.

We evaluated the segmentation and prediction accuracies of the AC and FC sizes. For the quantitative evaluation

of the segmentation radiographs, we employed intersection over union (IOU) and Dice similarity coefficient scores [19]. We calculated the IOU and Dice scores for every five regions and then averaged them over all regions. For the AC and FC size prediction accuracy, we employed the mean absolute error (MAE) and adjacent accuracy. The prediction of AC and FC sizes is an ordinal classification problem with an inherent order between classes. The MAE is a commonly used metric in ordinal classification. We calculated the MAE as average differences between the predicted and ground-truth class labels (0-9 for DELTA PF, 0-11 for Minima S, and 0-12 for MASTER SL). Adjacent accuracy is often used to evaluate ordinal classification algorithms in various research fields; however, it has been described using different terminologies [62], [63]. Adjacent accuracy is similar to the exact accuracy used in nominal classification problems, which allows predictions to adjacent classes within distance t to also be correct. We express the adjacent accuracy as $ACC^{\pm t}$ with distance t . $ACC^{\pm 0}$ is the exact accuracy that considers only predictions identical to the ground truth as correct. The accuracy within one size error, $ACC^{\pm 1}$, is commonly used for evaluating preoperative THA templating accuracy [58], [59].

B. SEGMENTATION

U-net was trained using a training dataset for segmentation. The best model was selected by evaluating the mean IOU score for the validation dataset at each epoch of the training process. Fig. 10 shows the changes in training loss and validation IOU scores according to model learning. Table 2 presents the IOU and Dice scores for the five regions. Small regions such as J and T had lower IOU and Dice scores. In deep learning-based segmentation, a small region is difficult because of an imbalanced problem [19], [20]. Additionally, the J and T regions had blurry boundaries in the original hip AP radiographs. Specifically, the T region often does not

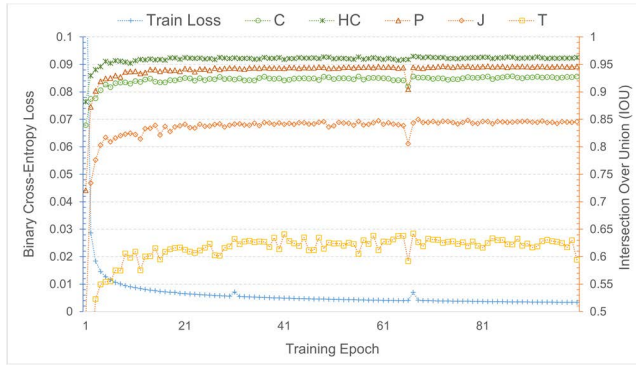


FIGURE 10. Curves of training loss and validation IOU scores during model training for region segmentation. The IOU scores are high in the order of HC, P, C, J, and T.

TABLE 2. IOU scores and dice coefficients for each of five regions for the train and validation sets.

Region	Training		Validation	
	IOU	Dice	IOU	Dice
C (Cortex)	0.970	0.985	0.932	0.965
HC (Head and Canal)	0.992	0.996	0.967	0.983
P (Pelvis)	0.988	0.994	0.949	0.974
J (Joint space)	0.965	0.982	0.850	0.919
T (Teardrop)	0.909	0.952	0.639	0.780
Mean	0.965	0.982	0.868	0.924

appear on radiographs, depending on the patient’s age and pose when acquiring the image.

We analyzed the cases of success and failure. Fig. 11 presents two successful cases with the highest IOU scores and two failure cases with the lowest IOU scores. Failures often occur when input images with different characteristics from the training data are used, such as bones appearing blurry or overlapping obstacles such as tubes in the original radiographic image. Fortunately, templating was correctly performed in most segmentation failure cases because the segmentation result identified most of the C, J, and femoral head (HC) regions.

C. ACETABULAR AND FEMORAL COMPONENT SIZE PREDICTIONS

We trained the CNN for AC size classification using the training dataset and evaluated the MAE using verification datasets at each epoch in the training process to select the best model. The training loss and validation accuracy curves are depicted in Fig. 12. We then evaluated the entire proposed procedure. Our templating program was designed to predict and place the most relevant THA components on AP hip radiographs. We evaluated the accuracy of predicting the sizes of the AC and FC using the test dataset. This experiment was applied to the healthy contralateral hip joint of the radiographs, according to the procedure shown in Fig. 2. As a result, the AC size prediction was generated for all 128 patients, whereas the FC was not generated and rejected for one patient because

of a significant matching error. Fig. 13 presents the confusion matrix of the prediction results for the AC and FC, and Table 3 shows the MAE and adjacent accuracies calculated from the confusion matrix.

The MAE for the AC and FC were 0.898 and 1.016, respectively. Most predictions, 94.1% for all components, were within two size differences. The exact accuracy $ACC^{\pm 0}$ was low at 35.2% and 37.0% for the AC and FC, respectively. However, the adjacent accuracy $ACC^{\pm 1}$ is acceptable at 78.9% for the AC and 70.9% for the FC. This result is reasonable compared with the meta-analysis of existing templating methods [58], [59]. As a result of the proposed method for all prostheses, $ACC^{\pm 0}$ was 36.1%, which is higher than 35% (min=25%; max=36%) of 2D acetate and lower than 48% (min=43%; max=61%) of 2D digital. Similarly, $ACC^{\pm 1}$ of the proposed method is 74.9%, which is higher than the 72% (min=45%; max=78%) of 2D acetate and lower than the 80% (min=75%; max=82%) of 2D digital by meta-analysis [59]. In particular, we obtained 78.9% and 70.9% $ACC^{\pm 1}$ for each of the AC and FC using uncemented prostheses, comparable to 73% and 74% of meta-analyses, respectively [58].

V. DISCUSSION

Obtaining a true pelvis AP radiography is clinically essential to evaluate anatomical landmarks as radiological ones. Radiologic landmarks can be changed by rotations and tilts of the pelvis or directions of the X-ray beam. In addition, obtaining a true AP image of the proximal femur is crucial because it has a wide morphological variation, including neck-shaft angle, offset, and torsion. Insufficient internal rotation of the lower extremity is one common cause of decreasing femoral offset. Therefore, the hip AP view should be obtained with an appropriate internal rotation to detect native femoral offset. It should be filmed by neutralizing the native femoral torsion and version to express the anatomical femoral offset well.

Difficulty in obtaining datasets is a limitation of this study. For the segmentation task, each region to be segmented must be labeled to train the U-net, which is tedious and time-consuming. To the best of our knowledge, there is no open- source hip joint segmentation data. Thus, we created 151 labeled image datasets for segmentation; however, they were small and insufficiently diverse. For this reason, segmentation may be unsatisfactory for an image with a characteristic different from the dataset, as shown in Fig. 11.

For AC size estimation, small and imbalanced datasets were used to train the CNN. The AC size clinically used is biased towards 50 and 54 mm. This is a characteristic of THA prostheses and its influence on surgical philosophy. THA surgeons tend to use a large ball head to improve the surgical outcome. Small and imbalanced datasets were improved by data augmentation and synthetic data generation.

FCs for prostheses differ in shape and size; ACs have similar shapes. It is difficult to construct a training dataset to classify different FCs. Although two types of FC families were used in the experiment, we aim to use the proposed

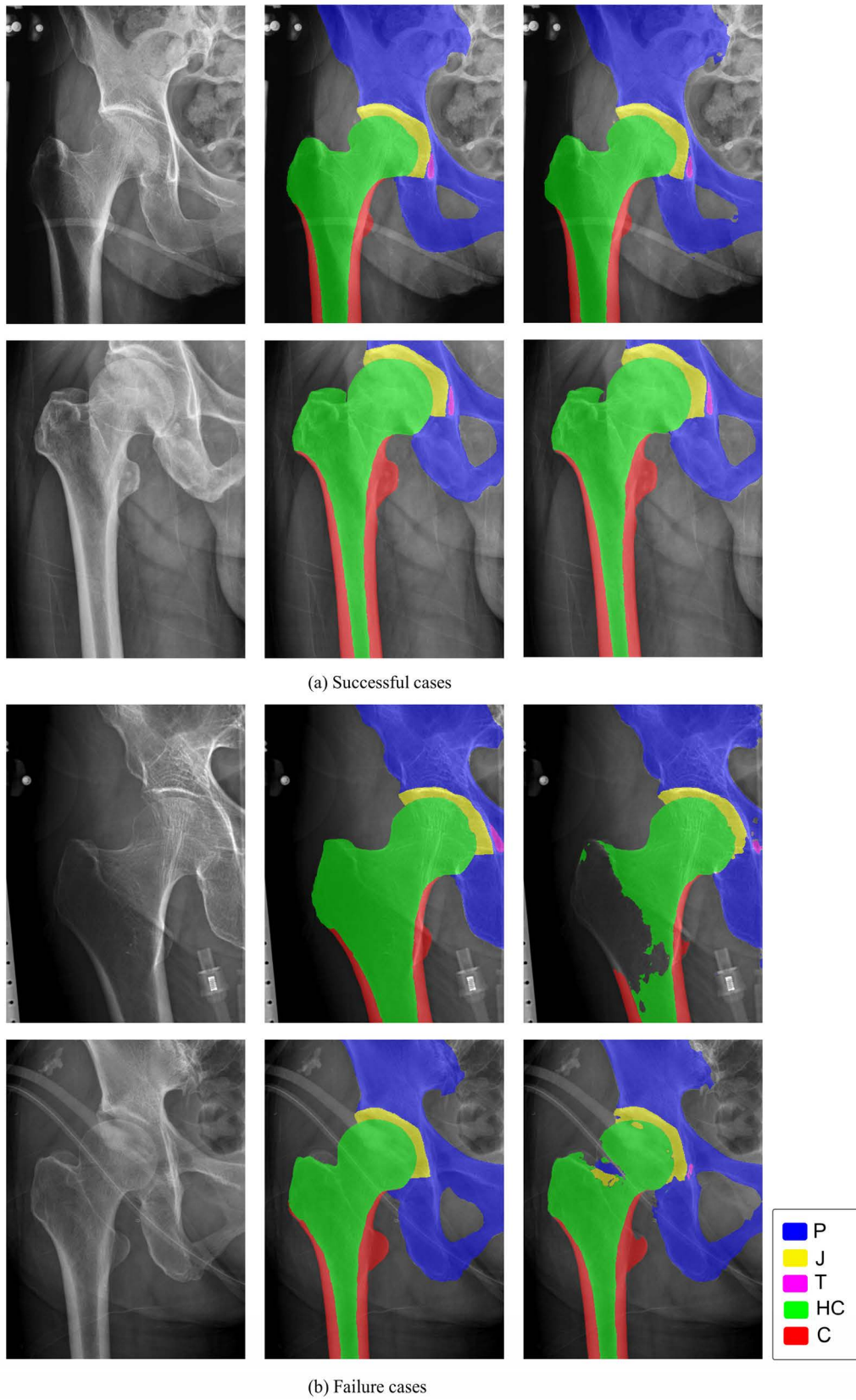


FIGURE 11. Segmentation results chosen concerning the IOU scores (left: input image patch, middle: ground-truth, right: segmentation result). Each colored region is as follows: P: Pelvis, T: Teardrop, J: Joint space and saucily, HC: Femoral head and canal, and C: Cortex.

TABLE 3. Performance of the acetabular and femoral component size estimation for the test dataset.

Method	Target	MAE	ACC±0	ACC±1	ACC±2
2D Automated (Proposed)	Acetabular DELTA PF	0.898	0.352	0.789	0.961
	Femoral Minima S	1.054	0.304	0.750	0.911
	MASTER SL	0.986	0.423	0.676	0.930
	Sub-total	1.016	0.370	0.709	0.921
	All components	0.957	0.361	0.749	0.941
2D Acetate [59]	All components	-	0.35	0.72	-
2D Digital [59]	All components	-	0.48	0.80	-
2D Digital [58]	Acetabular Uncemented	-	-	0.73	-
	Cemented	-	-	0.78	-
	Femoral Uncemented	-	-	0.74	-
	Cemented	-	-	0.89	-

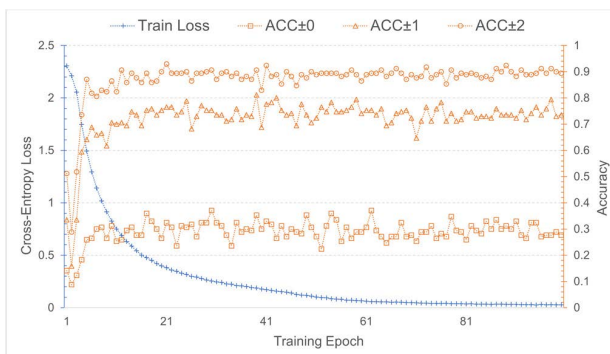


FIGURE 12. Curves of training loss and validation accuracies during model training to predict acetabular component size. ACC±0 is the exact accuracy, and ACC±1 and ACC±2 are the accuracies within one and two size errors, respectively.

method regardless of shape. In fact, the prototype system we developed includes 12 FC families. Thus, we first used a handcrafted algorithm to fit the FCs, and we will gradually replace them with learning-based methods in future research.

With segmentation, there was a difference in IOU scores according to the anatomical region. The highest IOU scores were observed in the diaphyseal cortex and femoral canal. This is consistent with actual physician insight. The T region, although an anatomically important structure, has individual variations, resulting in a low IOU score. Segmentation failure primarily occurs when there is little difference in opacity between soft tissue and bone, such as in osteoporotic bone quality or muscular patients. To improve performance, the radiation dose should be optimized to increase the difference in opacity, or the dataset should be expanded.

The accuracy was higher for the preoperative templating results using the deep learning-based prototype software developed by the authors than with previously used 2-D acetate methods. The results were consistent with those reported in recently published meta-analysis studies [58], [59]. The accuracy of the proposed method is acceptable considering its automation and completion within 30 s; the methods used in the meta-analysis are manual and require more time.

		Prediction										
		Size	44	46	48	50	52	54	56	58	60	62
Ground Truth	44	0	1	2	0	0	0	0	0	0	0	0
	46	2	1	2	1	0	0	0	0	0	0	0
	48	1	3	3	5	2	1	0	0	0	0	0
	50	2	2	9	19	9	9	1	0	0	0	0
	52	0	0	0	1	6	3	0	0	0	0	0
	54	0	0	1	2	10	15	6	0	0	0	0
	56	0	0	0	0	0	3	1	2	0	0	0
	58	0	0	0	0	0	2	0	0	0	0	0
	60	0	0	0	0	0	0	1	0	0	0	0
	62	0	0	0	0	0	0	0	0	0	0	0

(a) LimaCorporate DELTA PF cup

		Prediction												
		No.	1	2	3	4	5	6	7	8	9	10	11	12
Ground Truth	1	3	1	1	0	0	0	0	0	0	0	0	0	0
	2	0	0	2	1	1	0	0	0	0	0	0	0	0
	3	0	1	3	2	3	1	1	0	0	0	0	0	0
	4	1	1	3	3	3	1	0	0	0	0	0	0	0
	5	0	0	1	1	1	5	0	0	0	0	0	0	0
	6	0	0	1	0	2	2	1	0	0	0	0	0	0
	7	0	0	0	0	0	4	2	0	0	0	0	0	0
	8	0	0	0	0	0	0	0	2	0	0	1	0	0
	9	0	0	0	0	0	0	0	0	0	1	0	0	0
	10	0	0	0	0	0	0	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	0	0	0	0
	12	0	0	0	0	0	0	0	0	0	0	0	0	0

(b) LimaCorporate Minima S stem

		Prediction													
		No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Ground Truth	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	5	1	0	0	0	0	0	0	0	0	0	0	0
	3	0	1	2	0	2	0	0	0	0	0	0	0	0	0
	4	3	2	3	1	2	1	0	0	0	0	0	0	0	0
	5	0	0	3	1	5	1	1	0	0	0	0	0	0	0
	6	0	0	0	2	2	8	0	2	0	0	0	0	0	0
	7	0	0	1	0	0	4	2	1	0	0	0	0	0	0
	8	0	0	0	0	0	1	0	2	0	1	0	0	0	0
	9	0	0	0	0	0	0	2	1	3	0	0	0	0	0
	10	0	0	0	0	0	0	1	0	0	1	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	1	1	0	0	0
	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(c) LimaCorporate MASTER SL stem

FIGURE 13. Confusion matrix of the acetabular and femoral component size prediction for the test dataset.

Another advantage of this method is that it follows the thought process of a surgeon using radiologic landmarks used in actual hip surgery. This means that the deep

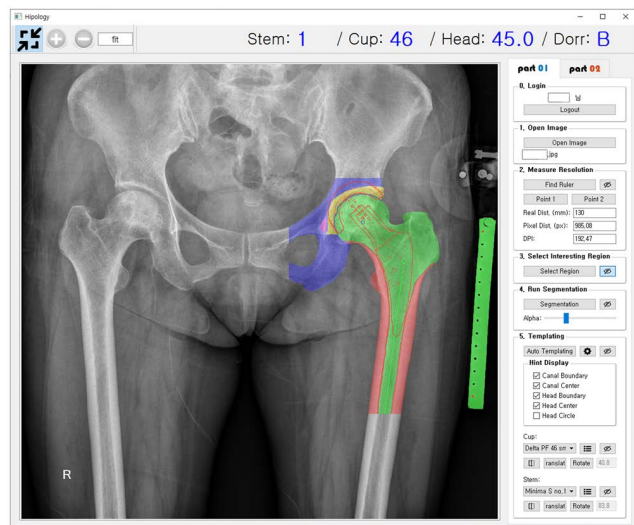


FIGURE 14. Preoperative THA planning prototype system.

learning-trained prototype software can be used in many different types of surgeries. For example, if the ROI is expanded, it can be applied to periacetabular osteotomy, intramedullary nailing, and fracture surgery. If the ROI is reduced and segmentation precision is enhanced, it can be applied to hip arthroscopy for femoroacetabular impingement.

Based on the programs developed in Section 3, we implemented a user interface and developed automated preoperative THA planning software, as shown in Fig. 14. The software provides functions for determining the ruler in the input image to measure the resolution, automatically predicting the THA component size in millimeters, and placing it in the proper position. Adjusting automatic templating results and manual templating is also possible. Additionally, the segmentation results can measure the femoral head size, sphericity, cortical thickness index, and canal-to-calcar ratio [64]. To evaluate its field adaptability, the software is being tested in an operation theater at Jeonbuk National University Hospital.

The main weaknesses of this work are the small and unbalanced datasets, rule-based algorithms that are easy to implement but struggle to provide above-human-level performances, and difficulties in positioning accuracy analysis. Based on the prototype system, our future research is intended to consider three aspects: (1) improving accuracy, (2) intra- and post-operative utilization, and (3) evaluation of clinical effectiveness. To improve accuracy, we plan to obtain more data and apply improved deep learning techniques. In addition, we will collect the logs of the planning results when orthopedic surgeons use our prototype system. These data can be used as training and validation data to predict the proper size and positioning of prostheses using deep learning, thereby replacing the rule-based methods. To increase the utilization of the proposed method, we consider using landmark detection and registration techniques to superimpose preoperative planning results on the intra-/post-operative images.

VI. CONCLUSION

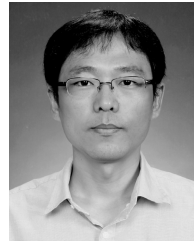
This study proposes an automated 2-D digital templating method for preoperative THA planning. The method consists of deep learning models for hip joint segmentation, AC size estimation, and handcrafted algorithms to fit the FC. Computer-aided surgery (CAS) software based on these methods has been developed and evaluated for field adaptability. The system widens the coverage of image analysis for AC and FC size estimation, and broadens applicability to all patients using simple inexpensively obtained radiographic images. The proposed data augmentation method is effective in preventing overfitting. The high segmentation and prediction accuracy of deep learning models can help advance preoperative structural templating into clinical practice for total hip arthroplasty. Critical future research will evaluate the clinical effectiveness of the CAS software and report it to the clinical community. Other research should focus on improving the system accuracy by using larger datasets and improved deep-learning techniques. Both are actively being conducted.

REFERENCES

- [1] D. Valle, A. González, D. E. Padgett, and E. A. Salvati, "Preoperative planning for primary total hip arthroplasty," *J. Amer. Acad. Orthopaedic Surg.*, vol. 13, no. 7, pp. 455–462, Nov. 2005.
- [2] A. Colombi, D. Schena, and C. C. Castelli, "Total hip arthroplasty planning," *EFORT Open Rev.*, vol. 4, no. 11, pp. 626–632, Nov. 2019, doi: 10.1302/2058-5241.4.180075.
- [3] J. V. Bono, "Digital templating in total hip arthroplasty," *J. Bone Joint Surg.*, vol. 86, pp. 118–122, Dec. 2004.
- [4] M. Moraliidou, A. Di Laura, J. Henckel, H. Hothi, and A. J. Hart, "Three-dimensional pre-operative planning of primary hip arthroplasty: A systematic literature review," *EFORT Open Rev.*, vol. 5, no. 12, pp. 845–855, Dec. 2020, doi: 10.1302/2058-5241.5.200046.
- [5] J. M. Vigdorich, A. K. Sharma, S. A. Jerabek, D. J. Mayman, and P. K. Sculco, "Templating for total hip arthroplasty in the modern age," *J. Amer. Acad. Orthopaedic Surgeons*, vol. 29, no. 5, pp. 208–216, Dec. 2020, doi: 10.5435/jaaos-d-20-00693.
- [6] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [7] E. Gibson, W. Li, C. Sudre, L. Fidon, D. I. Shaker, G. Wang, Z. Eaton-Rosen, R. Gray, T. Doel, Y. Hu, T. Whyntie, P. Nachev, M. Modat, D. C. Barratt, S. Ourselin, M. J. Cardoso, and T. Vercauteren, "NiftyNet: A deep-learning platform for medical imaging," *Comput. Methods Programs Biomed.*, vol. 158, pp. 113–122, May 2018, doi: 10.1016/j.cmpb.2018.01.025.
- [8] J. Ma, Y. Song, X. Tian, Y. Hua, R. Zhang, and J. Wu, "Survey on deep learning for pulmonary medical imaging," *Frontiers Med.*, vol. 14, no. 4, pp. 450–469, Aug. 2020, doi: 10.1007/s11684-019-0726-4.
- [9] X. Xie, J. Niu, X. Liu, Z. Chen, S. Tang, and S. Yu, "A survey on incorporating domain knowledge into deep learning for medical image analysis," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101985, doi: 10.1016/j.media.2021.101985.
- [10] E. Çallı, E. Sogancıoğlu, B. van Ginneken, K. G. van Leeuwen, and K. Murphy, "Deep learning for chest X-ray analysis: A survey," *Med. Image Anal.*, vol. 72, Aug. 2021, Art. no. 102125, doi: 10.1016/j.media.2021.102125.
- [11] F. Piccialli, V. D. Somma, F. Giampaolo, S. Cuomo, and G. Fortino, "A survey on deep learning in medicine: Why, how and when?" *Inf. Fusion*, vol. 66, pp. 111–137, Feb. 2021, doi: 10.1016/j.inffus.2020.09.006.
- [12] R. Szeliski, "Semantic segmentation," in *Computer Vision: Algorithms and Applications*, 2nd ed. Cham, Switzerland: Springer, 2022, ch.6, sec.4, pp. 307–315, doi: 10.1007/978-3-030-34372-9.

- [13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440, doi: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965).
- [14] X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 1089–1106, Aug. 2019, doi: [10.1007/s10462-018-9641-3](https://doi.org/10.1007/s10462-018-9641-3).
- [15] S. Ghosh, N. Das, I. Das, and U. Maulik, "Understanding deep learning techniques for image segmentation," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–35, Jul. 2020, doi: [10.1145/3329784](https://doi.org/10.1145/3329784).
- [16] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022, doi: [10.1109/TPAMI.2021.3059968](https://doi.org/10.1109/TPAMI.2021.3059968).
- [17] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *J. Digit. Imag.*, vol. 32, no. 4, pp. 582–596, Aug. 2019, doi: [10.1007/s10278-019-00227-x](https://doi.org/10.1007/s10278-019-00227-x).
- [18] J. Peng and Y. Wang, "Medical image segmentation with limited supervision: A review of deep network models," *IEEE Access*, vol. 9, pp. 36827–36851, 2021, doi: [10.1109/ACCESS.2021.3062380](https://doi.org/10.1109/ACCESS.2021.3062380).
- [19] M. Z. Khan, M. K. Gajendran, Y. Lee, and M. A. Khan, "Deep neural architectures for medical image semantic segmentation: Review," *IEEE Access*, vol. 9, pp. 83002–83024, 2021, doi: [10.1109/ACCESS.2021.3086530](https://doi.org/10.1109/ACCESS.2021.3086530).
- [20] S. A. Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and G. Hamarneh, "Deep semantic segmentation of natural and medical images: A review," *Artif. Intell. Rev.*, vol. 54, no. 1, pp. 137–178, Jan. 2021, doi: [10.1007/s10462-020-09854-1](https://doi.org/10.1007/s10462-020-09854-1).
- [21] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015, pp. 234–241, doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [22] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021, doi: [10.1109/ACCESS.2021.3086020](https://doi.org/10.1109/ACCESS.2021.3086020).
- [23] L. Joscowicz and E. J. Hazan, "Computer aided orthopaedic surgery: Incremental shift or paradigm change?" *Med. Image Anal.*, vol. 33, pp. 84–90, Oct. 2016, doi: [10.1016/j.media.2016.06.036](https://doi.org/10.1016/j.media.2016.06.036).
- [24] W. Li, I. Kezele, D. L. Collins, A. Zijdenbos, J. Keyak, J. Kornak, A. Koyama, I. Saeed, A. LeBlanc, T. Harris, Y. Lu, and T. Lang, "Voxel-based modeling and quantification of the proximal femur using inter-subject registration of quantitative CT images," *Bone*, vol. 41, no. 5, pp. 888–895, Nov. 2007, doi: [10.1016/j.bone.2007.07.006](https://doi.org/10.1016/j.bone.2007.07.006).
- [25] Y. Cheng, S. Zhou, Y. Wang, C. Guo, J. Bai, and S. Tamura, "Automatic segmentation technique for acetabulum and femoral head in CT images," *Pattern Recognit.*, vol. 46, no. 11, pp. 2969–2984, Nov. 2013, doi: [10.1016/j.patcog.2013.04.006](https://doi.org/10.1016/j.patcog.2013.04.006).
- [26] Y.-J. Yun, B.-C. Ahn, M. S. Kavitha, and S.-I. Chien, "An efficient region precise thresholding and direct Hough transform in femur and femoral neck segmentation using pelvis CT," *IEEE Access*, vol. 8, pp. 110048–110058, 2020, doi: [10.1109/ACCESS.2020.3001578](https://doi.org/10.1109/ACCESS.2020.3001578).
- [27] M. Krcak, G. Szekely, and R. Blanc, "Fully automatic and fast segmentation of the femur bone from 3D-CT images with no shape prior," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano Macro*, Mar. 2011, pp. 2087–2090, doi: [10.1109/ISBI.2011.5872823](https://doi.org/10.1109/ISBI.2011.5872823).
- [28] H. Liu, J. Zhao, N. Dai, H. Qian, and Y. Tang, "Improve accuracy for automatic acetabulum segmentation in CT images," *Bio-Med. Mater. Eng.*, vol. 24, no. 6, pp. 3159–3177, 2014.
- [29] Z. Zou, S.-H. Liao, S.-D. Luo, Q. Liu, and S.-J. Liu, "Semi-automatic segmentation of femur based on harmonic barrier," *Comput. Methods Programs Biomed.*, vol. 143, pp. 171–184, May 2017, doi: [10.1016/j.cmpb.2017.03.005](https://doi.org/10.1016/j.cmpb.2017.03.005).
- [30] F. Yokota, T. Okada, M. Takao, N. Sugano, Y. Tada, and Y. Sato, "Automated segmentation of the femur and pelvis from 3D CT data of diseased hip using hierarchical statistical shape model of joint structure," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2009, pp. 811–818, doi: [10.1007/978-3-642-04271-3_98](https://doi.org/10.1007/978-3-642-04271-3_98).
- [31] F. Yokota, T. Okada, M. Takao, N. Sugano, Y. Tada, N. Tomiyama, and Y. Sato, "Automated CT segmentation of diseased hip using hierarchical and conditional statistical shape models," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2013, pp. 190–197, doi: [10.1007/978-3-642-40763-5_24](https://doi.org/10.1007/978-3-642-40763-5_24).
- [32] D. Kainmueller, H. Lamecker, S. Zachow, and H.-C. Hege, "An articulated statistical shape model for accurate hip joint segmentation," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 6345–6351, doi: [10.1109/IEMBS.2009.5333269](https://doi.org/10.1109/IEMBS.2009.5333269).
- [33] J. Huang, J. F. Griffith, D. Wang, and L. Shi, "Graph-cut-based segmentation of proximal femur from computed tomography images with shape prior," *J. Med. Biol. Eng.*, vol. 35, no. 5, pp. 594–607, Oct. 2015, doi: [10.1007/s40846-015-0079-7](https://doi.org/10.1007/s40846-015-0079-7).
- [34] C. Chu, J. Bai, X. Wu, and G. Zheng, "MASC: Multi-atlas segmentation constrained graph method for accurate segmentation of hip CT images," *Med. Image Anal.*, vol. 26, no. 1, pp. 173–184, Dec. 2015, doi: [10.1016/j.media.2015.08.011](https://doi.org/10.1016/j.media.2015.08.011).
- [35] C. Chu, C. Chen, L. Liu, and G. Zheng, "FACTS: Fully automatic CT segmentation of a hip joint," *Ann. Biomed. Eng.*, vol. 43, no. 5, pp. 1247–1259, May 2015, doi: [10.1007/s10439-014-1176-4](https://doi.org/10.1007/s10439-014-1176-4).
- [36] Y. Chang, Y. Yuan, C. Guo, Y. Wang, Y. Cheng, and S. Tamura, "Accurate pelvis and femur segmentation in hip CT with a novel patch-based refinement," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 3, pp. 1192–1204, May 2019, doi: [10.1109/JBHI.2018.2834551](https://doi.org/10.1109/JBHI.2018.2834551).
- [37] G. Zeng, X. Yang, J. Li, L. Yu, P. A. Heng, and G. Zheng, "3D U-Net with multi-level deep supervision: Fully automatic segmentation of proximal femur in 3D MR images," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2017, pp. 274–282.
- [38] C. M. Deniz, S. Xiang, R. S. Hallyburton, A. Welbeck, J. S. Babb, S. Honig, K. Cho, and G. Chang, "Segmentation of the proximal femur from MR images using deep convolutional neural networks," *Sci. Rep.*, vol. 8, no. 1, p. 16485, Nov. 2018, doi: [10.1038/s41598-018-34817-6](https://doi.org/10.1038/s41598-018-34817-6).
- [39] F. Chen, J. Liu, Z. Zhao, M. Zhu, and H. Liao, "Three-dimensional feature-enhanced network for automatic femur segmentation," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 1, pp. 243–252, Jan. 2019, doi: [10.1109/JBHI.2017.2785389](https://doi.org/10.1109/JBHI.2017.2785389).
- [40] C. Lindner, S. Thiagarajah, J. M. Wilkinson, G. A. Wallis, and T. F. Cootes, "Accurate fully automatic femur segmentation in pelvic radiographs using regression voting," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2012, pp. 353–360.
- [41] S. Schumann, L. Liu, M. Tannast, M. Bergmann, L.-P. Nolte, and G. Zheng, "An integrated system for 3D hip joint reconstruction from 2D X-rays: A preliminary validation study," *Ann. Biomed. Eng.*, vol. 41, no. 10, pp. 2077–2087, May 2013, doi: [10.1007/s10439-013-0822-6](https://doi.org/10.1007/s10439-013-0822-6).
- [42] C. Chen, W. Xie, J. Franke, P. A. Grützner, L. P. Nolte, and G. Zheng, "Automatic X-ray landmark detection and shape segmentation via data-driven joint estimation of image displacements," *Med. image Anal.*, vol. 18, no. 3, pp. 487–499, Apr. 2014, doi: [10.1016/j.media.2014.01.002](https://doi.org/10.1016/j.media.2014.01.002).
- [43] W. Xie, J. Franke, C. Chen, P. A. Grützner, S. Schumann, L.-P. Nolte, and G. Zheng, "Statistical model-based segmentation of the proximal femur in digital antero-posterior (AP) pelvic radiographs," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 9, no. 2, pp. 165–176, Mar. 2014, doi: [10.1007/s11548-013-0932-5](https://doi.org/10.1007/s11548-013-0932-5).
- [44] R. Pilgram, C. Walch, V. Kuhn, R. Schubert, and R. Staudinger, "Proximal femur segmentation in conventional pelvic X ray," *Med. Phys.*, vol. 35, no. 6, pp. 2463–2472, May 2008, doi: [10.1118/1.2919096](https://doi.org/10.1118/1.2919096).
- [45] F. Ouertani, C. Vazquez, T. Cresson, and J. de Guise, "Simultaneous extraction of two adjacent bony structures in X-ray images: Application to hip joint segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4555–4559, doi: [10.1109/ICIP.2015.7351669](https://doi.org/10.1109/ICIP.2015.7351669).
- [46] P. Gamage, S. Q. Xie, P. Delmas, and W. L. Xu, "Segmentation of radiographic images under topological constraints: Application to the femur," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 5, no. 5, pp. 425–435, Jan. 2010.
- [47] W. Shen, W. Xu, H. Zhang, Z. Sun, J. Ma, X. Ma, S. Zhou, S. Guo, and Y. Wang, "Automatic segmentation of the femur and tibia bones from X-ray images based on pure dilated residual U-Net," *Inverse Problems Imag.*, vol. 15, no. 6, pp. 1333–1346, Dec. 2021, doi: [10.3934/ipi.2020057](https://doi.org/10.3934/ipi.2020057).
- [48] P. Rouzrokh, C. C. Wyles, K. A. Philbrick, T. Ramazanian, A. D. Weston, J. C. Cai, M. J. Taunton, D. G. Lewallen, D. J. Berry, B. J. Erickson, and H. M. Kremers, "A deep learning tool for automated radiographic measurement of acetabular component inclination and version after total hip arthroplasty," *J. Arthroplasty*, vol. 36, no. 7, pp. 2510–2517, Jul. 2021, doi: [10.1016/j.arth.2021.02.026](https://doi.org/10.1016/j.arth.2021.02.026).
- [49] E. Schiffner, D. Latz, P. Jungbluth, J. P. Grassmann, S. Tanner, A. Karbowski, J. Windolf, and J. Schneppendahl, "Is computerised 3D templating more accurate than 2D templating to predict size of components in primary total hip arthroplasty?" *HIP Int.*, vol. 29, no. 3, pp. 270–275, May 2019, doi: [10.1177/1120700018776311](https://doi.org/10.1177/1120700018776311).

- [50] J. Huo, G. Huang, D. Han, X. Wang, Y. Bu, Y. Chen, D. Cai, and C. Zhao, "Value of 3D preoperative planning for primary total hip arthroplasty based on artificial intelligence technology," *J. Orthopaedic Surg. Res.*, vol. 16, no. 1, p. 156, Dec. 2021, doi: [10.1186/s13018-021-02294-9](https://doi.org/10.1186/s13018-021-02294-9).
- [51] M. Brenneis, S. Braun, S. van Drongelen, B. Fey, T. Tarhan, F. Stief, and A. Meurer, "Accuracy of preoperative templating in total hip arthroplasty with special focus on stem morphology: A randomized comparison between common digital and three-dimensional planning using biplanar radiographs," *J. Arthroplasty*, vol. 36, no. 3, pp. 1149–1155, Mar. 2021, doi: [10.1016/j.arth.2020.10.016](https://doi.org/10.1016/j.arth.2020.10.016).
- [52] P. Gamble, J. de Beer, D. Petrucci, and M. Winemaker, "The accuracy of digital templating in uncemented total hip arthroplasty," *J. Arthroplasty*, vol. 25, no. 4, pp. 529–532, Jun. 2010, doi: [10.1016/j.arth.2009.04.011](https://doi.org/10.1016/j.arth.2009.04.011).
- [53] L. A. Holzer, G. Scholler, S. Wagner, J. Friesenbichler, W. Maurer-Ertl, and A. Leithner, "The accuracy of digital templating in uncemented total hip arthroplasty," *Arch. Orthopaedic Trauma Surg.*, vol. 139, no. 2, pp. 263–268, Feb. 2019, doi: [10.1007/s00402-018-3080-0](https://doi.org/10.1007/s00402-018-3080-0).
- [54] X. Ding, B. Zhang, W. Li, J. Huo, S. Liu, T. Wu, and Y. Han, "Value of preoperative three-dimensional planning software (AI-HIP) in primary total hip arthroplasty: A retrospective study," *J. Int. Med. Res.*, vol. 49, no. 11, Nov. 2021, Art. no. 0300060521105888, doi: [10.1177/03000605211058874](https://doi.org/10.1177/03000605211058874).
- [55] D. Dammerer, A. Keiler, S. Herrnegger, D. Putzer, S. Strasser, and M. Liebensteiner, "Accuracy of digital templating of uncemented total hip arthroplasty at a certified arthroplasty center: A retrospective comparative study," *Arch. Orthopaedic Trauma Surg.*, pp. 1–10, Mar. 2021, doi: [10.1007/s00402-021-03836-w](https://doi.org/10.1007/s00402-021-03836-w).
- [56] E. Kristoffersson, V. Otten, and S. Crnalic, "The accuracy of digital templating in cementless total hip arthroplasty in dysplastic hips," *BMC Musculoskeletal Disorders*, vol. 22, no. 1, p. 942, Dec. 2021, doi: [10.1186/s12891-021-04793-6](https://doi.org/10.1186/s12891-021-04793-6).
- [57] L. T. Buller, A. S. McLawhorn, J. D. Maratt, K. M. Carroll, and D. J. Mayman, "EOS imaging is accurate and reproducible for preoperative total hip arthroplasty templating," *J. Arthroplasty*, vol. 36, no. 3, pp. 1143–1148, Mar. 2021, doi: [10.1016/j.arth.2020.09.051](https://doi.org/10.1016/j.arth.2020.09.051).
- [58] J. B. V. Smith, H. Bishi, C. Wang, V. Asopa, R. E. Field, and D. H. Sochart, "The accuracy and reliability of preoperative digital 2D templating in prosthesis size prediction in uncemented versus cemented total hip arthroplasty: A systematic review and meta-analysis," *EFORT Open Rev.*, vol. 6, no. 11, pp. 1020–1039, Nov. 2021, doi: [10.1302/2058-5241.6.210048](https://doi.org/10.1302/2058-5241.6.210048).
- [59] H. Bishi, J. B. V. Smith, V. Asopa, R. E. Field, C. Wang, and D. H. Sochart, "Comparison of the accuracy of 2D and 3D templating methods for planning primary total hip replacement: A systematic review and meta-analysis," *EFORT Open Rev.*, vol. 7, no. 1, pp. 70–83, Jan. 2022, doi: [10.1530/EOR-21-0060](https://doi.org/10.1530/EOR-21-0060).
- [60] I. Otomaru, K. Kobayashi, T. Okada, M. Nakamoto, Y. Kagiya, M. Takao, N. Sugano, Y. Tada, and Y. Sato, "Expertise modeling for automated planning of acetabular cup in total hip arthroplasty using combined bone and implant statistical atlases," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2009, pp. 532–539, doi: [10.1007/978-3-642-04268-3_66](https://doi.org/10.1007/978-3-642-04268-3_66).
- [61] P. F. Felzenszwalb and D. P. Huttenlocher, "Distance transforms of sampled functions," *Theory Comput.*, vol. 8, no. 1, pp. 415–428, 2012, doi: [10.4086/toc.2012.v008a019](https://doi.org/10.4086/toc.2012.v008a019).
- [62] Y. Peng, W. Zhu, Z. Chen, M. Wang, L. Geng, K. Yu, Y. Zhou, T. Wang, D. Xiang, F. Chen, and X. Chen, "Automatic staging for retinopathy of prematurity with deep feature fusion and ordinal classification strategy," *IEEE Trans. Med. Imag.*, vol. 40, no. 7, pp. 1750–1762, Jul. 2021, doi: [10.1109/TMI.2021.3065753](https://doi.org/10.1109/TMI.2021.3065753).
- [63] E.-G. Kim, I.-S. Oh, J.-E. So, J. Kang, V. N. T. Le, M.-K. Tak, and D.-W. Lee, "Estimating cervical vertebral maturation with a lateral cephalogram using the convolutional neural network," *J. Clin. Med.*, vol. 10, no. 22, p. 5400, Nov. 2021, doi: [10.3390/jcm10225400](https://doi.org/10.3390/jcm10225400).
- [64] S.-J. Yoon, M. Kim, I.-S. Oh, K. Kim, and K.-S. Han, "Estimation and comparison of cortical thickness index and canal-to-calcar ratio using manual method and deep learning method," *J. Electr. Eng. Technol.*, vol. 15, no. 3, pp. 1399–1404, May 2020, doi: [10.1007/s42835-020-00387-9](https://doi.org/10.1007/s42835-020-00387-9).



MINWOO KIM received the B.S., M.S., and Ph.D. degrees in computer engineering from Jeonbuk National University, Jeonju, South Korea, in 2007, 2009, and 2016, respectively. From 2013 to 2019, he was a Researcher and a Developer with the Research and Development Center, Gachisoft Inc., and ST-1 Company Ltd., South Korea. Since 2019, he has been a Research Professor with the Division of Computer Science and Engineering, Jeonbuk National University. His

research interests include computer vision, pattern recognition, and artificial intelligence.



IL-SEOK OH received the B.S. degree in computer engineering from Seoul National University, South Korea, in 1984, and the Ph.D. degree in computer science from the KAIST, South Korea, in 1992. He was a Visiting Scientist with the CEN-PARMI, Concordia University, Canada, and the UCI, USA. He is currently a Professor with the Division of Computer Science and Engineering, Jeonbuk National University, Jeonju, South Korea. He is the author of the books *Pattern Recognition*,

Computer Vision, and *Machine Learning* (written in Korean). His research interests include computer vision, pattern recognition, and machine learning.



SUN-JUNG YOON received the M.D. degree in medicine, the Ph.D. degree in tissue engineering and biomaterial science, and the Ph.D. degree in orthopedic surgery from Jeonbuk National University, Jeonju, South Korea, in 2003 and 2010, respectively. He is currently working as a Chief Surgeon of Hip and Pelvis Service with the Department of Orthopedic Surgery, Jeonbuk National University Hospital, Jeonju. His research interests

include adult hip reconstruction, hip preservation surgery, artificial intelligence, biomaterial science, and drug delivery systems. He was a recipient of the Yuhan Young Investigator Award (the Korean Medical Association), the Jeonbuk National University Hospital Research Award, and the Best Research Award of the *Hip & Pelvis* journal (the Korean Hip Society).

• • •