BMC Musculoskeletal Disorders

RESEARCH

Open Access

Deep learning-based automated measurement of hip key angles and auxiliary diagnosis of developmental dysplasia of the hip

Ruixin Li[1,2†], Xiao Wang[1†], Tianran Li[1*], Beibei Zhang[1], Xiaoming Liu[3], Wenhua Li[1] and Qirui Sui[1]

Abstract

Objectives Anteroposterior pelvic radiographs remains the most widely employed method for diagnosing developmental dysplasia of the hip. This study aims to evaluate the accuracy of an artificial intelligence model in measuring angles in pelvic radiographs of the hip. The assessment seeks to demonstrate the efficacy of the artificial intelligence model in diagnosing both developmental dysplasia of the hip and borderline developmental dysplasia of the hip through the analysis of pelvic radiographs.

Methods A total of 1,029 patients, including 273 men and 757 women, were retrospectively included in this study. The anteroposterior pelvic radiographs were randomly divided into three sets: the training set (720 cases), the validation set (103 cases), and the test set (206 cases). Key anatomical points on the anteroposterior pelvic radiographs were identified. The Sharp, Tönnis, and Center Edge angles were calculated automatically based on the corresponding criteria. The hip development status was compared between measurements obtained from the artificial intelligence model and those defined manually by two radiologists. The area under the receiver operating characteristic curve was utilized to assess the diagnostic performance of the artificial intelligence model.

Results The results obtained from both manual measurements and the artificial intelligence model demonstrated no significant differences in the Sharp, Tönnis, and Center edge angles (all $p > 0.05$). The intra-class correlation coefficients and correlation coefficient r values exceeded 0.75, indicating that both the artificial intelligence model and manual measurements exhibited good repeatability and a positive correlation. Notably, the artificial intelligence model provided measurements more faster than those conducted by radiologists ($p = 0.001$). The artificial intelligence model also demonstrated high diagnostic accuracy, sensitivity, and specificity for developmental dysplasia of the hip. The performance of the artificial intelligence model in diagnosing developmental dysplasia of the hip was robust. Additionally, the results from the artificial intelligence model and manual measurements were largely consistent with clinical diagnosis results ($p = 0.01$). The artificial intelligence model can effectively evaluate hip conditions by measuring the Sharp, Tönnis, and Center edge angles, which are consistent closely with clinical diagnosis results.

Conclusions The results of the artificial intelligence model measurements demonstrate a high degree of consistency with those obtained through manual measurements. The angles of Sharp, Tönnis, and Center edge, as evaluated by the deep learning-based convolutional neural network model, exhibit robust diagnostic performance in identifying both developmental dysplasia of the hip and borderline developmental dysplasia of the hip. Consequently,

†Ruixin Li and Xiao Wang are the co-first authors.

*Correspondence:
Tianran Li
ltranmd@yeah.net
Full list of author information is available at the end of the article

the artificial intelligence model has the potential to fully replace manual measurements of these critical hip angles, providing a more efficient and precise alternative for diagnosing both conditions of the hip.

Keywords Developmental dysplasia of the hip, Artificial intelligence, Deep learning, Pelvic radiograph

Background

Developmental dysplasia of the hip (DDH) is one of the most prevalent bone and joint disorders in the field of orthopedics [1]. DDH encompasses a range of pathologies affecting the acetabulum and proximal femur, including acetabular dysplasia, hip subluxation, true hip dislocation, and hip instability [2]. Additionally, borderline acetabular dysplasia (BDDH) refers to mild abnormalities in acetabular shape and coverage that may lead to mechanical dysfunction and instability [3]. Due to the long treatment cycles and a relatively high disability rates, some patients even require lifelong follow-up and treatment. Therefore, early diagnosis and early treatment are the most effective methods to improve the survival status of patients with DDH [4, 5]. However, early DDH is often asymptomatic or presents with mild symptoms, and there are no significant changes in medical imaging, leading to many cases being missed or misdiagnosed [6]. Untreated DDH can ultimately lead to the onset of secondary osteoarthritis [7]. Epidemiological studies indicate that the incidence of DDH ranges from 5 to 30% [8]. Radiographic examination remains the most widely employed method for the diagnosis of DDH and BDDH [3, 9]. The anteroposterior pelvic radiographs is one of the most commonly used techniques for hip examination in clinical practice. It plays a crucial role in disease assessment, preoperative planning, and postoperative efficacy evaluation [10, 11]. This requires the accurate measurement of the main parameters of the hip joint to achieve accurate diagnosis or classification [12–14]. Various measurement methods from different angles and dimensions have been proposed, among which the most widely recognized imaging measurement parameters in clinical practice include the Sharp angle, Center Edge angle (CE angle), and Tönnis angle [15]. However, manual measurements obtained from anteroposterior pelvic X-rays exhibit considerable errors and poor reproducibility. This is primarily due to the strong subjectivity involved in different observers' identification of anatomical landmarks, often relying on clinical experience, which leads to poor diagnostic consistency [16, 17]. Consequently, many patients miss the optimal timing for early diagnosis and treatment [18]. Therefore, there is an urgent clinical need for an intelligent tool that can measure anteroposterior pelvic radiographs in a large-scale, objective, simple, and precise methods, while efficiently and quickly screening for suspected DDH or BDDH patients. The ability to accurately measure hip joint angles is crucial for influencing early clinical diagnosis and treatment decisions.

In recent years, artificial intelligence (AI) has achieved remarkable advancements in image recognition, segmentation, decision-making, and the quantitative analysis of extensive datasets [18–20]. In certain domains, it has even outperformed clinicians [21]. Deep learning, as a representative technology, has been applied to the auxiliary management of hip joint diseases [22]. Fengyu et al. [23] utilized a single-shot multi-box detector to achieve hip joint detection on anteroposterior pelvic radiographs. However, their study focused solely on the localization of the hip joint and did not provide any evidence to support the diagnosis of DDH. In another study [24], the left-right symmetry of 11 key points in the hip joint was utilized, with the interconnected edges serving as anatomical constraints to identify the anatomical positions of multiple key points. This approach enabled the calculation of indicators necessary for diagnosing various hip joint diseases. However, due to the broad scope of the research objectives, the findings of this study do not offer targeted clinical guidance or facilitate decision-making for specific diseases.

We therefore posed the following questions: (1) How can we more accurately and efficiently measure the angles of the hip for the diagnosis of DDH and BDDH? (2) What is the level of consistency and repeatability between angles measured manually and those obtained through an AI model on anteroposterior pelvic radiographs of the hip? (3) How effective is the AI model, based on anteroposterior hip pelvic X-ray radiographs, in assisting with the diagnosis of DDH and BDDH?

As a result, it is to improve the accuracy of early diagnosis of DDH, reduce errors in manual measurement of hip joint angles by radiologists, and enhance efficiency, this study employs deep learning algorithms to identify key anatomical points on anteroposterior pelvic radiographs. A convolutional neural network AI model based on deep learning algorithms is constructed for the automatic measurement of the Sharp angle, CE angle, and Tönnis angle. Additionally, the effectiveness of the AI automatic measurement model of the key angle of the hip joint in assisting the diagnosis of DDH and BDDH is evaluated, providing support for the clinical application of AI in assessing DDH.

Methods

General and patient information

A total of 1,029 patients' data, consisting of 273 men and 757 women (aged 18–84 years, median age: 33 years), who underwent anteroposterior pelvic radiographs examinations at the hospital between January 2020 and January 2022, were retrospectively included in this study. The original anteroposterior pelvic radiography images were retrieved, and the data were processed using the picture archiving and communication system (PACS).

Criteria for inclusion: (1) Each anatomical key point is clearly delineated; (2) Patients are at least 18 years of age with closed epiphyses; (3) Anteroposterior pelvic radiographs meet diagnostic quality standards as confirmed by experienced radiologists.

Exclusion criteria: (1) Presence of flat hip malformation or a history of pelvic surgery; (2) Hip joint abnormalities resulting from rheumatism, infection, or neoplastic diseases; (3) Secondary dislocation of the hip joint due to central nervous system disorders or trauma; (4) Indistinguishable tear drop or excessive proliferation of the acetabulum; (5) Poor X-ray positioning or insufficient image data.

The study employed clinical diagnostic criteria—comprising clinical presentation, clinical examination, radiographs, and ultrasound screening—as the gold standard [1, 5].

Data sets

According to the results of our previous study, there was no statistically significant difference in the measurement of angles when comparing pelvic X-ray radiographs of patients with DDH taken in standing and supine positions [15]. Consequently, anteroposterior pelvic radiographs obtained in both supine and standing positions were utilized to create an image database. The images were randomly divided into three sets: the training set (720 cases, accounting for 70% of the total), the validation set (103 cases, accounting for 10% of the total), and the test set (206 cases, accounting for 20% of the total). The random grouping of images was assigned using a random number table method. The training set was employed to develop the artificial intelligence model for identifying key anatomical points of the hip joint, while the validation set was used to optimize the model parameters. The test set served to evaluate and compare the accuracy of key point localization, angle measurements, and diagnostic performance for DDH and BDDH (Fig. 1).

Instrument and X-ray radiography methods

The Definium 6000 digital X-ray radiography system (GE Medical Imaging Systems, USA) was utilized for the X-ray examination of the hip. The X-ray examination was conducted as follows: (1) In the standard supine position, the patient was instructed to lie on their back with the tips of both first toes turned inward at an angle of 15° and in contact with each other; the lower limbs were to be kept straight and shoulder-width apart. The central line was defined as the line passing through the midpoint between the upper edge of the pubic symphysis and the bilateral anterior superior iliac spines. (2) In the standard standing position, the patient was instructed to stand in front of the camera with legs straight and both feet slightly turned inward at an angle of 15°. The central line was established through a point located 2–3 cm above the pubic symphysis.

Measurement parameters and joint marking

The most commonly utilized and widely recognized hip joint measurement parameters, specifically the

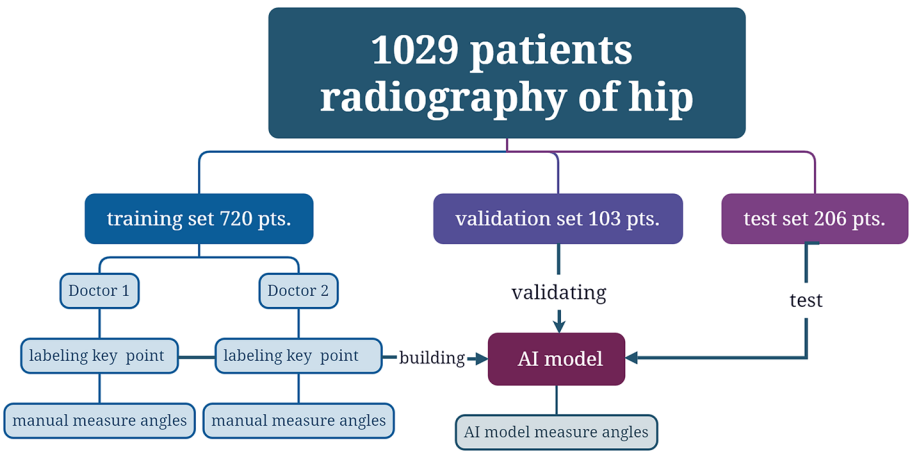

Fig. 1 Illustrates the patient grouping and flowchart

Sharp, CE, and Tönnis angles [15], were selected for this study (Fig. 2). These parameters were delineated by radiologists with 5 years (Dr. 1) and 3 years (Dr. 2) of relevant experience. The labeled anteroposterior pelvic radiographs re-checked by a senior radiologist, were subsequently included in the training set.

The following anatomical points were identified: the center point of the left femoral head (L-fhc), the point on the outer edge of the left acetabulum (L-uar), the point on the inner edge of the left acetabulum (L-tar), the point on the lower edge of the left tear drop (L-lt), the center point of the right femoral head (R-fhc), the point on the outer edge of the right acetabulum (R-uar), the point on the inner edge of the right acetabulum (R-tar), and the point on the lower edge of the right tear drop (R-lt) (Fig. 3). Points exhibiting discrepancies greater than 5 mm were re-evaluated and marked again by the two radiologists and a senior radiologist to ensure the objectivity of the marking process.

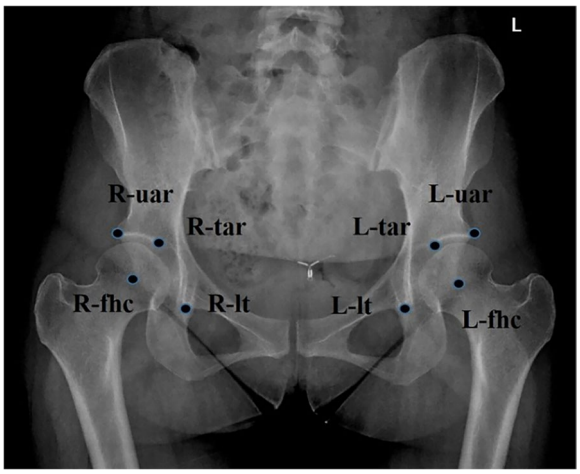

Fig. 3 Schematic representation of the hip joint marking. Select center point of the left femoral head as L-fhc, and point on the outer edge of the left acetabulum as L-uar, and point on the inner edge of the left acetabulum as L-tar, and point on the lower edge of the left tear drop as L-lt, and center point of the right femoral head as R-fhc, and point on the outer edge of the right acetabulum as R-uar, and point on the inner edge of the right acetabulum as R-tar, and point on the lower edge of the right tear drop as R-lt

Hip angle measurement and auxiliary diagnosis of DDH and BDDH

The Sharp, Tönnis, and CE angles were calculated automatically based on the a forementioned coordinates and corresponding criteria. Diagnosis of DDH or BDDH was established according to the following conditions [25]: (1) DDH was diagnosed when the Sharp angle exceeded 45°; (2) DDH was also diagnosed when the Tönnis angle was greater than 10°; (3) BDDH was diagnosed when the CE angle ranged from 20° to 25° [26, 27], whereas DDH was diagnosed if the CE angle measured less than 20°.

Data preprocessing

To enhance the dataset size, we collected data from various acquisition devices across the training, validation, and test sets, and performed separate preprocessing to mitigate discrepancies, thereby improving the robustness of subsequent models. The pixel dimensions of all radiographs were standardized to 0.2 mm × 0.2 mm through B-spline interpolation, a data resampling technique. Subsequently, the pixel values were normalized to a unified range using min-max normalization.

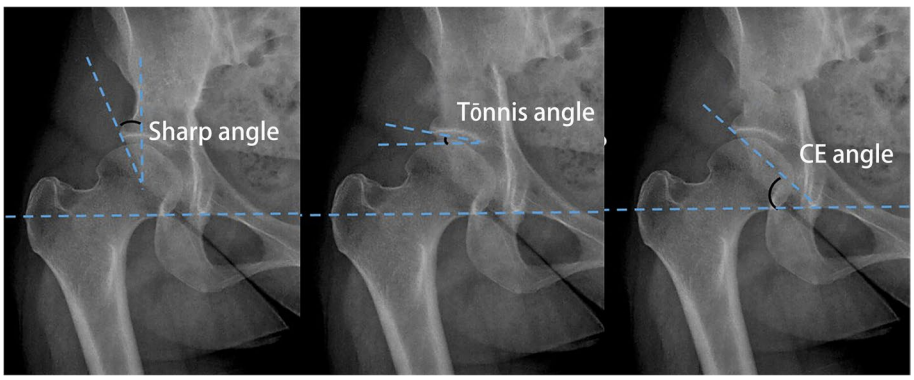

Fig. 2 Schematic representation of the measurement of the CE, Tönnis, and Sharp angles. **A** CE angle measurement method, the angle between the perpendicular line from the center point of the femoral head and the outer edge of the acetabulum. **B** Tönnis angle measurement method, the angle between the line connecting the lowest edge of the ilium and the outermost edge of the acetabulum and the H line (the H line, also known as the Hilgenreiner line, is the line connecting the tear drop points). **C** Sharp angle measurement method, the angle formed by the bilateral tear drop lines and the line connecting the lower end of the tear drop to the outer upper edge of the acetabulum

Finally, as illustrated in Fig. 4, all included images were flipped along the perpendicular bisector of the corresponding X-ray images. The left-sided images were mirrored to create the input dataset for locating key points of the hip anatomy. This approach reduced the number of targets for hip key point localization from 8 to 4, while simultaneously increasing the sample size and decreasing the task's complexity. During the training phase, the input radiographs were randomly rotated between $-5°$ and $5°$ to further enhance the model's robustness [28].

Model construction

In this study, we employed VB-Net (Visual Basic Net) as the foundational network for locating key points of the hip anatomical position [29]. Its architecture is illustrated in Fig. 5. The VB-Net innovatively replaces the convolutional layers of the traditional U-Net architecture with a bottleneck structure, significantly compressing network parameters while preserving model performance and enhancing training efficiency. The bottleneck block comprises three convolutional layers: the first layer utilizes a $1\times1\times1$ convolution kernel to reduce the channel dimensions of the feature maps; the second layer conducts

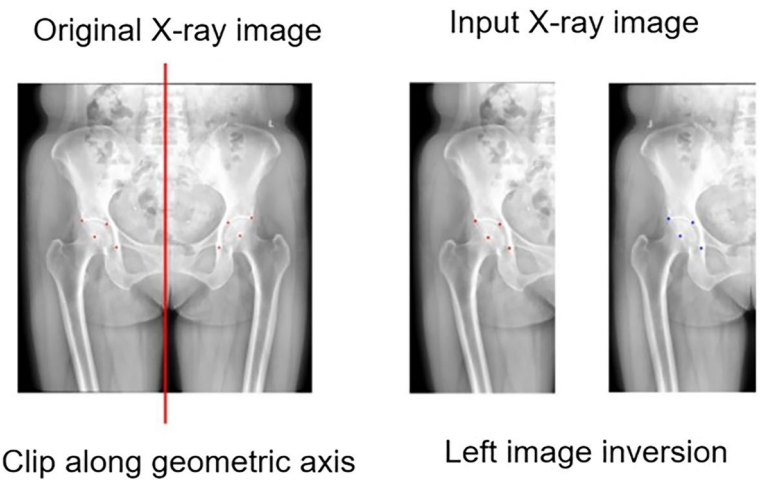

Fig. 4 Schematic representation of data preprocessing

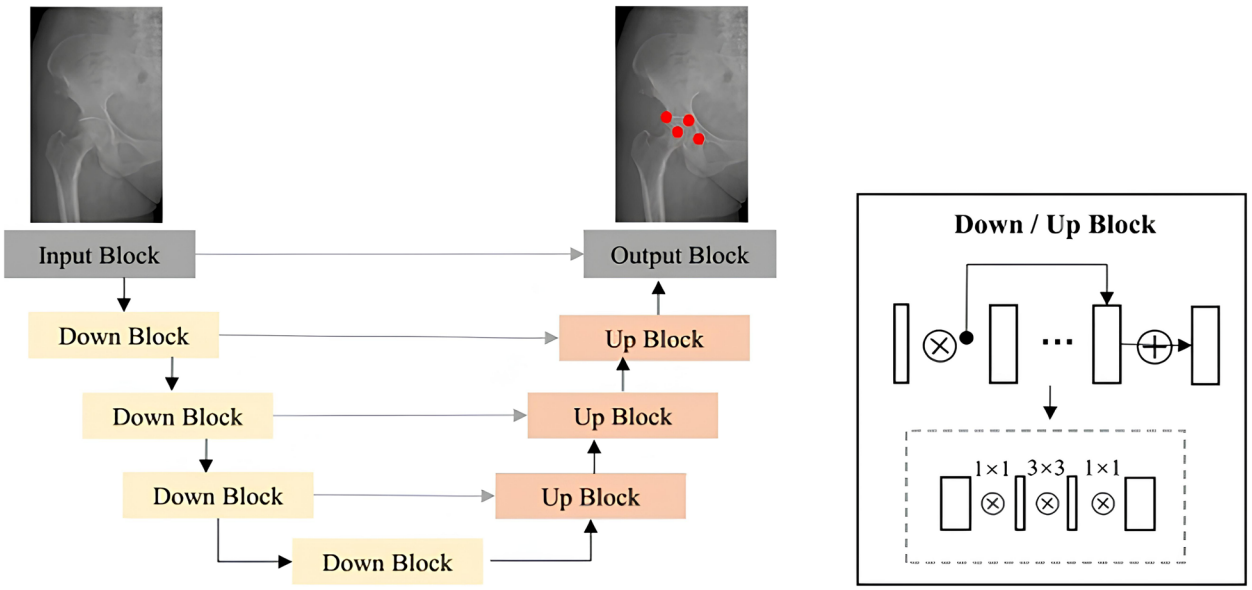

Fig. 5 A schematic representation of the automated measurement of hip key points utilizing the VB-Net model

spatial convolution with a kernel size of 3×3×3 to effectively encode the input; and the final layer again applies a 1×1×1 convolution kernel to restore the channel dimensions of the feature maps to their original size. During model training, we defined the ground truth as a patch region of 20 pixels × 20 pixels surrounding each anatomical landmark for input into VB-Net. In the inference phase, we identified the center of the largest connected component in the probability maps, applying a threshold of 0.5. This approach facilitated landmark detection through a segmentation network. By integrating spatial information, the coordinates of the predicted points from the left-sided image were mapped to the original image coordinate system. Subsequently, the Sharp, Tönnis, and CE angles of the left and right hips were calculated in the test set according to the aforementioned definitions.

In this study, focal loss was employed as the loss function for the network, with the learning rate set to 0.0001. Data preprocessing and automated parameter calculations were conducted using Python 3.8. The training and evaluation of the deep learning model were carried out utilizing the PyTorch framework on an NVIDIA GeForce GTX 1080Ti graphics card. A validation set of 103 cases was used for further optimization of the model parameters in the test set, employing a cross-validation method.

Comparison between AI automatic measurement and manual measurement of hip joint angles

Two radiologists manually measured the Sharp angle, Tönnis angle, and CE angle of the test set (206 cases) using tools provided by PACS, taking the average value of each radiologist's measurements as the statistical value. If there is a significant discrepancy between the measurements of the key hip joint angles by the two radiologists, a senior radiologist will conduct a data re-checking.

Based on the above rules, an AI automatic measurement of the Sharp angle, Tönnis angle, and CE angle was performed on a test set of 206 cases. Firstly, we compared the differences between the AI model automatic measurement and manual measurements (MM) of key angles of the hip joint, which serves to evaluate the accuracy of the AI model measurement results. Secondly, a comparison was made regarding the time taken for the AI model to measure the key angles versus manual measurements of the hip joint. Furthermore, the diagnostic performance of the AI model in assessing DDH was evaluated, along with a comparison of the diagnostic efficacy of the AI model and radiologists for DDH and BDDH, with diagnoses made according to the criteria [25–27]. Finally, using clinical diagnosis results as the "gold standard," the results of the AI model and manual diagnosis of DDH and BDDH were compared with the clinical diagnosis results, and the purpose is to evaluate whether AI can be used in clinical practice.

Statistical analysis

Statistical analysis was performed using SPSS 25.0 software (IBM, USA). For continuous data (Sharp angle, Tönnis angle, and CE angle values, AI model measurements, and specific names listed for the time taken for manual measurement of key angles), a preliminary normality test was conducted. Data that followed a normal distribution were expressed as mean ± standard deviation. The comparison of angle measurements between Dr.1, Dr.2, and the AI model was analyzed using the least significant difference (LSD) method. The correlation between AI model measurements and the manual measurements by radiologists (Dr.1, Dr.2) was assessed using Pearson correlation analysis. The consistency of measurement results was evaluated using the intraclass correlation coefficient (ICC), with an ICC ≥ 0.75 indicating good consistency. The diagnostic efficacy of the AI model was assessed using the area under the receiver operating characteristic curve (ROC curve, AUC). For categorical data (the number of cases for AI, radiologists, and the final clinical diagnoses of DDH and BDDH), the consistency of AI and manual measurement results with the final clinical diagnosis was evaluated using the Kappa test. The significance level (α) was set at two-sided 0.05.

Results

Evaluation of the accuracy of AI measurement key angles results

An AI automatic measurement model for hip joint key angles was constructed using a training dataset of 720 cases through deep learning. The model parameters were optimized using a validation dataset of 103 cases. Currently, the model's measurements of the Sharp angle, Tönnis angle, and CE angle are compared with MM using a test dataset of 206 anteroposterior pelvic radiographs to evaluate the accuracy of the AI model's measurement results and the corresponding findings are presented in Table 1.

There were no statistically significant differences between the manually measured and AI model-measured results for Sharp, Tönnis, and CE angles (all $p > 0.05$).

The evaluation of consistency between AI model measurements and MM for the Sharp angle, Tönnis angle, and CE angle is presented in Fig. 6. The ICC values for the left side Sharp angle, Tönnis angle, and CE angle are 0.801, 0.757, and 0.895, respectively. For the right side, the ICC values for the Sharp angle, Tönnis angle, and CE angle are 0.943, 0.790, and 0.908, respectively. All ICC values are greater than 0.75, indicating

Table 1 Difference analysis between manual and AI model measurements (test set, $n=206$)

		Manual measurement		AI Model measurement	F	P
		Dr. 1	Dr. 2			
Left	Sharp angle	43.4°±5.2°	43.2°±5.3°	43.2°±5.9°	0.03	0.99
	Tönnis angle	12.2°±7.8°	11.9°±7.8°	12.3°±9.7°	0.08	0.97
	Center edge angle (CE)	25.4°±10.1°	25.4°±10.2°	25.6°±9.6°	0.05	0.98
Right	Sharp angle	43.5°±5.6°	43.1°±5.5°	43.0°±5.4°	0.36	0.78
	Tönnis angle	11.6°±8.5°	11.1°±8.1°	11.8°±9.5°	0.30	0.83
	Center edge angle (CE)	25.0°±10.2°	24.7°±10.4°	26.5°±9.9°	1.36	0.26

Dr. Doctor, *AI* Artificial intelligence

that the AI model measurements and MM of the hip joint key angles exhibit good consistency.

The Pearson correlation analysis between AI model measurements and MM for the Sharp angle, Tönnis angle, and CE angle shows that the r values for the left side are 0.807, 0.774, and 0.896, respectively, while the r values for the right side are 0.946, 0.794, and 0.921. This indicates that the AI model measurements and MM of the hip joint key angles have a high degree of correlation and a positive relationship.

It can be observed from Fig. 6 that all scatter points are situated within the standard deviation lines, and the mean line is approximately at zero. This indicates that the results obtained from the two measurement methods (manual measurement and AI model measurement) are relatively close and exhibit good consistency. Further analysis revealed that the mean absolute errors of the AI model measurements for the Sharp, Tönnis, and CE angles were 1.33°, 2.30°, and 2.68° on the left side, and 1.12°, 2.35°, and 3.05° on the right side, respectively. The root mean square errors for the AI model measurements of the Sharp, Tönnis, and CE angles were 3.49°, 6.13°, and 4.48° on the left side, and 1.85°, 5.86°, and 4.36° on the right side, respectively. These results suggest that the AI model measurements are highly consistent with the manual measurements, exhibiting smaller errors.

Comparison of time taken for hip measurement

The AI model took an average of 1.7 ± 0.00 s to measure key angles in 206 test set images, while Radiologist 1 (Dr. 1) took an average of 88.13 ± 8.41 s, and Radiologist 2 (Dr. 2) took an average of 90.28 ± 7.40 s (Fig. 7).

Radiologists required a longer duration to perform hip joint key angles measurements compared to the AI model measurements. Consequently, the AI model's measurement process was significantly faster than that of the radiologists ($p=0.000$).

Evaluation of the diagnostic performance of DDH through AI model measurements

The results of the AI-based automated measurement of the hip joint key angles, derived from anteroposterior pelvic radiographs, were analyzed. Three parameters utilized to assess hip development were employed for the auxiliary diagnosis of DDH (see Table 2). The findings indicated that all hip joint key angles measured by the AI model demonstrated high diagnostic accuracy, sensitivity, and specificity for DDH.

Comparison of the efficacy of AI measurement and manual measurement of hip joint angles for diagnosing DDH and BDDH

The receiver operating characteristic (ROC) curves for both the AI model and MM were utilized to evaluate the DDH and BDDH (Fig. 8).

The analysis presented in Fig. 8A and B indicates that the area under the curve (AUC) of the ROC for the AI model measuring the Sharp angle (left: 0.883, right: 0.924) and the Tönnis angle (left: 0.908, right: 0.922) closely aligns with that of experienced radiologists. This finding suggests that the AI model demonstrates high diagnostic performance for DDH. Furthermore, as illustrated in Fig. 8C, the CE angle measurements obtained by the AI model also exhibit significant diagnostic efficacy for DDH, with AUC values of 0.922 for the left side and 0.871 for the right side, The AI model demonstrates a high diagnostic efficacy for DDH.

Additionally, the CE angle measurements by the AI model provide valuable diagnostic support for BDDH, as evidenced by the AUC values of 0.787 for the left side and 0.676 for the right side, also depicted in Fig. 8C. Results indicate that AI model has a similarly high diagnostic efficacy.

Comparison of AI model and MM with clinical diagnosis results of DDH and BDDH

Using clinical diagnosis results as the gold standard, we compared the Sharp angle (AI model A), Tönnis angle

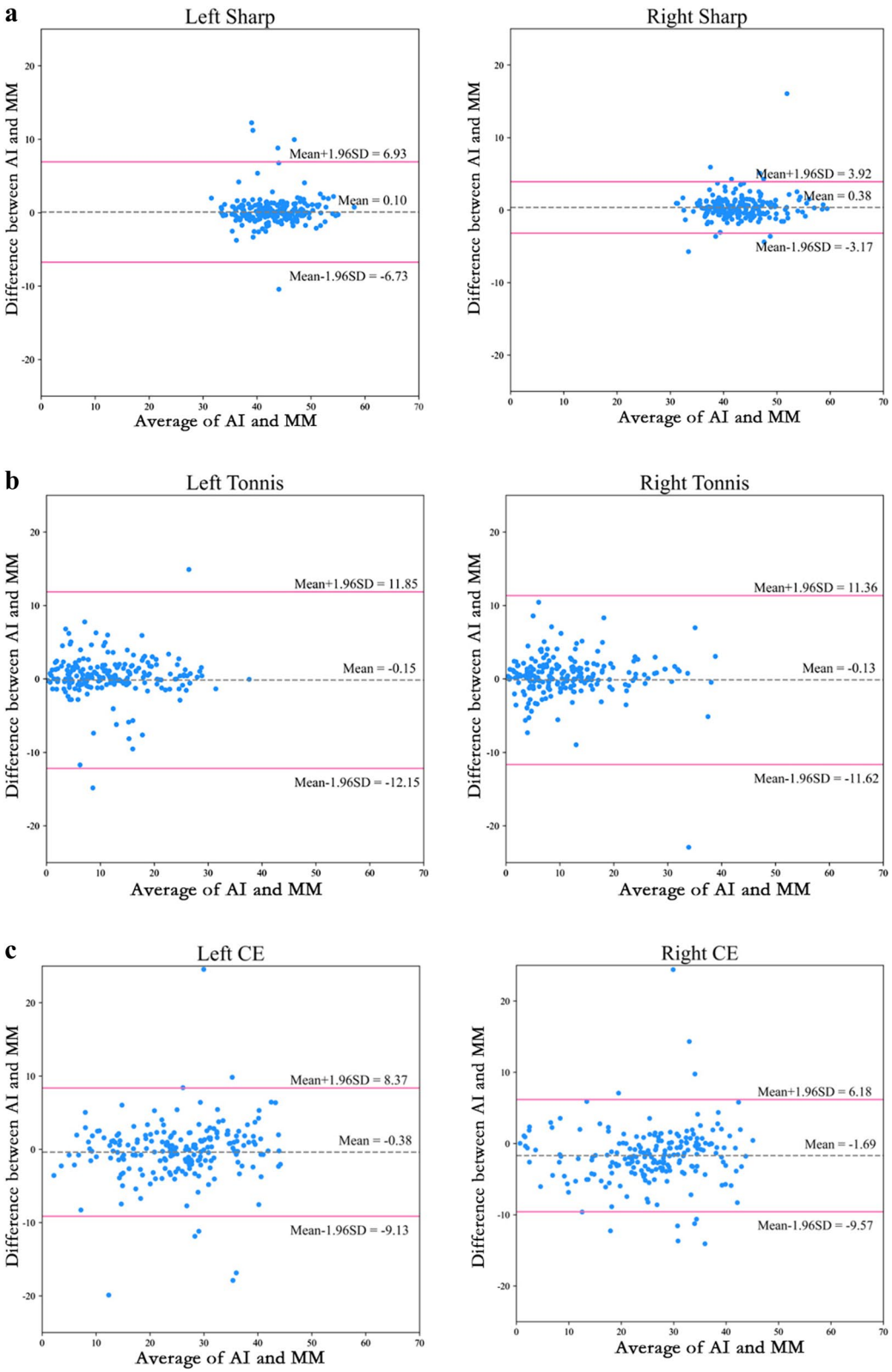

Fig. 6 The Bland-Altman plot illustrates the agreement between measurements obtained from the artificial intelligence model (AI) and those derived from manual measurement (MM) across various angles

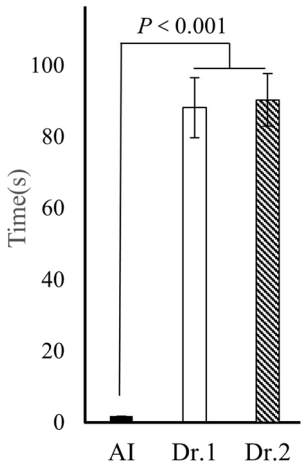

Fig. 7 The average time taken by the artificial intelligence (AI) model and two radiologists to perform measurements on each image

Table 2 Diagnostic performance of DDH based on artificial intelligence model measurement

Hip	Angle	Accuracy	Sensitivity	Specificity
Left	Sharp	89.8%	82.1%	94.5%
	Tönnis	90.1%	90.3%	91.4%
	Center edge (CE)	86.8%	84.8%	87.9%
Right	Sharp	93.7%	88.6%	96.3%
	Tönnis	92.2%	91.3%	93.2%
	Center edge (CE)	80.5%	90.9%	76.4%

(AI model B), and CE angle (AI model C) with the final clinical diagnosis outcomes, alongside manual measurement results. Kappa tests were conducted (see Tables 3, 4 and 5).

The results of the kappa tests indicated that the AI model and MM for the Sharp, Tönnis, and CE angles were largely consistent with the clinical diagnosis results ($p=0.000$). This finding suggests that the AI model can effectively evaluate hip conditions by measuring the Sharp, Tönnis, and CE angles, which align closely with clinical diagnosis outcomes, thereby serving as a valuable tool for the auxiliary diagnosis of DDH and BDDH.

Discussion

The hip joint, one of the largest load-bearing joints in the human body, primarily functions to provide dynamic support for the weight of the upper body while facilitating the axial transmission of stress and load to the lower limbs [30]. Its ball-and-socket structure offers significant horizontal stability and permits movement across three planes: flexion/extension, abduction/adduction, and internal/external rotation [31, 32]. Research has demonstrated that persistent hip dysplasia can lead to alterations in body posture and gait; furthermore, it suggests that patients with diminished cartilage or other degenerative joint diseases exhibit compensatory changes, often manifesting as limb shortening. These adaptations can result in postural scoliosis, back pain, or even disability, significantly impacting the quality of life for affected individuals. The Sharp, Tönnis, and CE angles serve as critical indicators for the clinical assessment of hip joint stability, enabling the evaluation of acetabular development and its coverage of the femoral head [33].

The measurement of the sharp angle of the hip joint, Tönnis angle, and CE angle has always been a cumbersome task for clinicians and radiologists. The determination of anatomical landmarks is often a prerequisite for precise angle measurement, which is highly subjective and leads to variations in the parameters measured manually, both within and between observers. This variability affects the accuracy of manual measurements. Delegating these repetitive clinical tasks to intelligent tools can significantly alleviate the workload of physicians and improve work efficiency. As a basis for measuring the key angles of the hip joint, this study utilizes the bilateral femoral head center points (flc), bilateral lower edge points of the teardrop (lt), bilateral inner edge points of the acetabulum (tar), and bilateral outer edge points of the acetabulum (uar) as key anatomical landmarks. The selection of these key points is consistent with the anatomical points commonly used in clinical measurements of hip joint data.

Some scholars have investigated the differences in measurement angles of pelvic X-ray radiographs taken in standing and supine positions [34]. However, our previous study indicated no statistically significant differences in angle measurements when pelvic X-ray radiographs of patients with DDH were obtained in either position [15]. Consequently, anteroposterior pelvic radiographs acquired in both supine and standing positions were utilized to create an image database.

Previous research has reported measurement errors among different observers ranging from ±3.5° to ±10° [19]. Yinjun et al. [35] employed a deep learning algorithm for the automated measurement of the Sharp angle of the hip, finding mean absolute errors of 1.57° and 1.73° for the left and right sides, respectively. In contrast, our results demonstrated mean absolute errors of 1.33° and 1.12° for the left and right Sharp angles, respectively, which were lower than those reported by Yinjun et al. [35]. Furthermore, we observed no significant differences between the angles measured by the AI model and those measured manually for the Sharp, Tennis, and CE angles. However, the AI model yielded smaller measurements for the Sharp angles on both the left and right sides.

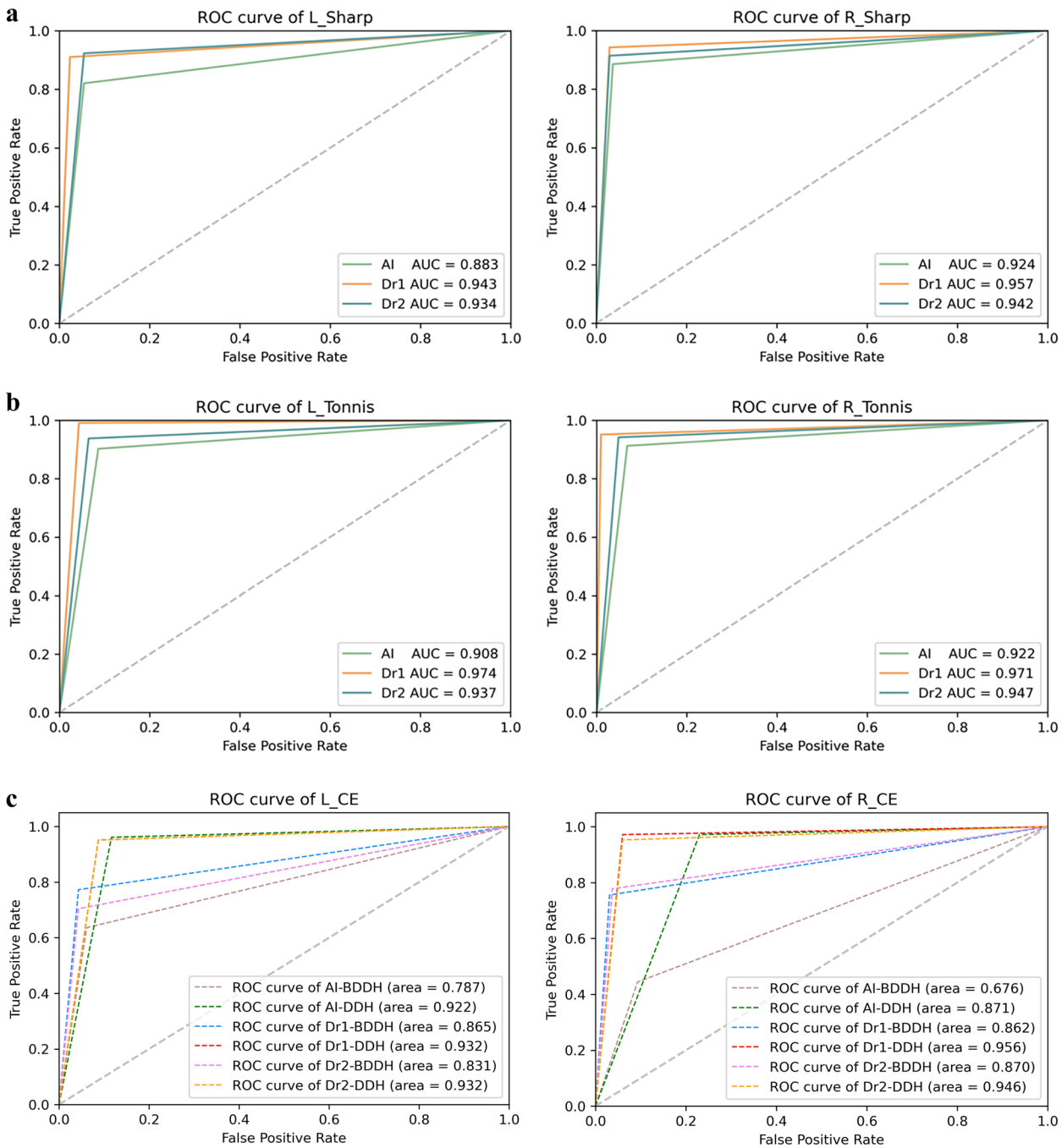

Fig. 8 ROC curves were generated based on the AI model and MM of the hip Sharp, Tönnis and CE angles for the diagnosis of DDH and BDDH

The presence of osteophytes at the outer edge of the acetabulum can influence the magnitude of the measured hip angles. Consequently, the magnitude of the measured CE angle may be significantly elevated if osteophytes are not clearly identified [36]. However, since none of the cases included in this study exhibited obvious osteophytes at the outer edge of the acetabulum, this factor is unlikely to have substantially impacted the results. Qiang et al. [37]. developed a model for the automated measurement of the CE angle, and the results produced by this model did not significantly differ from those measured by radiologists. Holden et al. [38] measured the Sharp, Tönnis, and CE angles after constructing a model that automatically identified key

Table 3 Comparison of AI model A (Sharp angle) and manual measurement diagnosis results with clinical diagnosis results ($n=206$)

	Diagnosis results	Diagnosis results							
		Left				Right			
		Normal	DDH	Kappa	P	Normal	DDH	Kappa	P
AI model A	Normal	121	14	0.779	0.000	131	8	0.858	0.000
	DDH	7	64			5	62		
Dr.1	Normal	125	7	0.898	0.000	133	4	0.913	0.000
	DDH	3	71			4	65		
Dr.2	Normal	121	6	0.860	0.000	133	6	0.890	0.000
	DDH	7	72			4	63		

DDH Developmental dysplasia of the hip, *Dr D*octor, *Normal* Normal hip

Table 4 Comparison of AI model B (Tönnis angle) and manual measurement diagnosis results with clinical diagnosis results ($n=206$)

	Diagnosis results	Diagnosis results							
		Left				Right			
		Normal	DDH	Kappa	P	Normal	DDH	Kappa	P
AI model B	Normal	85	11	0.814	0.000	96	9	0.845	0.000
	DDH	8	102			7	94		
Dr.1	Normal	92	1	0.951	0.000	102	5	0.942	0.000
	DDH	4	109			1	98		
Dr.2	Normal	87	7	0.873	0.000	98	6	0.893	0.000
	DDH	6	106			5	97		

DDH Developmental dysplasia of the hip, *Dr* Doctor, *Normal* Normal hip

Table 5 Comparison of AI model C (CE angle) and manual measurement diagnosis results with clinical diagnosis results ($n=206$)

	Diagnosis results	Diagnosis results									
		Left					Right				
		Normal	BDDH	DDH	Kappa	P	Normal	BDDH	DDH	Kappa	P
AI mode C	Normal	51	5	1	0.778	0.000	42	2	1	0.663	0.000
	BDDH	3	28	7			13	20	2		
	DDH	1	11	99			0	23	103		
Dr.1	Normal	98	9	0	0.866	0.000	103	6	0	0.873	0.000
	BDDH	5	34	2			3	34	2		
	DDH	0	1	57			0	5	53		
Dr.2	Normal	98	9	0	0.835	0.000	103	6	1	0.840	0.000
	BDDH	4	32	3			3	34	3		
	DDH	1	4	55			2	5	49		

DDH Developmental dysplasia of the hip, *Dr* Doctor, *Normal* Normal hip, *BDDH* Borderline developmental dysplasia of the hip

points of the hip joint. They reported a good consistency between the results obtained manually and those measured by the model. We replicated this finding and further used for demonstrated the feasibility of the AI model for diagnosing DDH and BDDH.

In the previous studies conducted by Qiang et al. [37], the constructed models required an average of 1.19 and 1.22 s to automatically measure the Sharp and CE angles, respectively. In our study, the constructed model took an average of 1.7 s to simultaneously measure the Sharp,

Tönnis, and CE angles, while two radiologists took an average of 88.13 ± 8.41 s (Dr.1) and 90.28 ± 7.40 s (Dr.2) to measure the three angles. Therefore, the AI model significantly enhances the efficiency of radiologists in performing these measurements. Compared to Li's model, our model incorporates the measurement of both Tönnis and CE angles, requiring a similar amount of time while improving diagnostic accuracy by adding testing methods and providing greater potential for clinical application.

Despite ongoing controversies regarding the diagnosis of BDDH [39], the classification of femoral head and acetabulum coverage can be delineated into BDDH and DDH based on varying degrees of morphological changes, which are characterized by an exceptionally shallow or upwardly inclined acetabulum. These alterations can result in uneven stress distribution across the hip joint [23]. Consequently, hips with CE angles of less than 20° or within the range of 20°–25° (or 18°–24°) are classified as having DDH or BDDH [3]. In this study, the diagnostic efficacy, as measured by the area under the curve (AUC), for the left and right CE angles in the artificial intelligence (AI) model for DDH was 0.922 and 0.871, respectively, while for BDDH, the AUC values were 0.787 and 0.676, respectively. These findings indicate that measurements derived from the AI model hold significant value in the auxiliary diagnosis of both DDH and BDDH. Hips exhibiting a Sharp angle greater than 45° and a Tönnis angle exceeding 10° were identified as dysplastic, with diagnostic accuracies of 93.7% and 92.2% for the left side, and 89.8% and 90.1% for the right side, respectively. Our research results indicate that the AI model measuring key angles of the hip joint has a high diagnostic efficacy in assisting the diagnosis of DDH and BDDH. Finally, using the clinical final diagnosis as the "gold standard," we observed the consistency between the AI model and clinical diagnosis, and the results showed a high level of agreement, fully meeting clinical needs.

It has been suggested that the prevalence of DDH is ten times lower in East Asian populations compared to White populations [40]. However, another study indicates that the incidence of DDH in East Asia appears to be comparable to that in Western countries [41]. Conducting a global multicenter study would significantly influence clinical decision-making regarding DDH.

This study has several limitations. First, the identification of key anatomical points on anteroposterior pelvic radiographs was performed manually, relying on the subjective experience of the operators. Consequently, slight deviations may occur between different operators. This lack of objective standard values results in a disparity between the diagnostic accuracy of the AI model and manual measurements. Second, external validation was not conducted; the data processing was limited to our hospital. Therefore, the generalizability of the model may be constrained. Although the model developed in this study can measure the Sharp, Tönnis, and CE angles, it can only provide a single indicator for evaluation, which poses significant challenges for its future application in clinical practice. Furthermore, to enhance diagnostic and treatment screening, further research is needed to develop measurement models for the femoral neck stem angle, acetabular coverage, and acetabular depth. Establishing CT and MRI data models and collaboratively building a multimodal intelligent diagnosis and treatment system for DDH will be essential to improve the detection rate of DDH.

Conclusions

In this study, a deep learning algorithm was employed to identify key anatomical points on anteroposterior pelvic radiographs and to develop a convolutional neural network model for the automated measurement of Sharp, Tönnis, and CE angles. The efficacy of the AI model in assisting the diagnosis of DDH was also assessed. The measurement results obtained from the AI model demonstrated a high degree of consistency with those derived from manual measurements, and both methods exhibited comparable repeatability. Notably, the AI model outperformed radiologists in terms of measurement accuracy. The angles of Sharp, CE, and Tönnis measured by the deep learning-based convolutional neural network model can be utilized for the diagnosis of DDH and BDDH with high precision. The AI model serves as a valuable tool to assist radiologists in measuring critical hip angles and diagnosing DDH and BDDH more swiftly and accurately.

Abbreviations

DDH	Developmental Dysplasia of the Hip
BDDH	Borderline Acetabular Dysplasia
AI	Artificial Intelligence
PACS	Picture Archiving and Communication System
L-fhc	Center Point of the Left Femoral Head
L-uar	Point on the Outer Edge of the Left Acetabulum
L-tar	Point on the Inner Edge of the Left Acetabulum
L-lt	Point on the Lower Edge of the Left Tear Drop
R-fhc	Center Point of the Right Femoral Head
R-uar	Point on the Outer Edge of the Right Acetabulum
R-tar	Point on the Inner Edge of the Right Acetabulum
R-lt	Point on the Lower Edge of the Right Tear Drop
CE	Center Edge
VB-Net	Visual Basic Net
ROC	The Receiver Operating Characteristic Curves

Acknowledgements

We would like to express our gratitude to the Department of Research and Development at United Imaging Intelligence (Beijing) Co., Ltd. for their support, particularly for the contributions of Dr. Xiaoming Liu. TianRan Li had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors have read and approved the final version of the manuscript.

Authors' contributions
RXL, XX, BBZ and TRL wrote the main manuscript text, and RXL, XX,XML prepared all Figs. RXL, XX and WHL participated in the data collection, and analysis. WX and TRL contributed to the study design and conception, and revision of the manuscript. All authors reviewed the manuscript.

Funding
This study was not supported by any research fund.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
The study received approval from the Institutional Review Board of the Fourth Medical Center of the General Hospital of the Chinese People's Liberation Army. All methods were conducted in accordance with relevant guidelines and regulations. The procedures adhered to the principles outlined in the Helsinki Declaration. We confirm that informed consent was obtained from all participants. Each author certifies that all investigations were carried out in compliance with ethical standards.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Radiology, The Fourth Medical Center of Chinese PLA General Hospital, No. 51, Fucheng Road, Haidian District, Beijing, China. [2]Department of General Surgery, The Fourth Medical Center of Chinese PLA General Hospital, No. 51, Fucheng Road, Haidian District, Beijing,, China. [3]Department of Research and Development, United Imaging Intelligence(Beijing) Co.Ltd, No. 9, Yongteng Road, Haidian District, Beijing, China.

Received: 1 May 2024 Accepted: 6 November 2024
Published online: 13 November 2024

References

1. Sioutis S, Kolovos S, Papakonstantinou ME, et al. Developmental dysplasia of the hip: a review. J Long Term Eff Med Implants. 2022;32:39–56.
2. Anand SD, Muzammil A, Daniel R, et al. Current surgical techniques in the treatment of adult developmental dysplasia of the hip. J Pers Med. 2023;13:942.
3. Sarah DB, Michael BM. The borderline dysplastic hip: when and how is it abnormal? Pediatr Radiol. 2019;49:1669–77.
4. Shaw BA, Segal LS, Section on Orthopaedics. Evaluation and referral for developmental dysplasia of the hip in infants. Pediatrics. 2016;138:e20163107.
5. Roof AC, Jinguji TM, White KK. Musculoskeletal screening: developmental dysplasia of the hip. Pediatr Ann. 2013;42:229–35.
6. Bridget KE, Ernest LS, Shevaun MD. Adolescent hip dysplasia: what are the symptoms and how to diagnose it. Curr Opin Pediatr. 2021;33:65–73.
7. Zhang S, Doudoulakis KJ, Khurwal A, et al. Developmental dysplasia of the hip. Br J Hosp Med. 2020;81:1–8.
8. Schmitz MR, Murtha AS, Clohisy JC, et al. Developmental dysplasia of the hip in adolescents and young adults. J Am Acad Orthop Surg. 2020;28:91–101.
9. Weize X, Liqi S, Ping G, et al. A deep learning aided diagnostic system in assessing developmental dysplasia of the hip on pediatric pelvic radiographs. Front Pediatr. 2022;9:785480.
10. Li X, Qu Y, Wang L, Ai S. The accuracy of cup anteversion measurement on postoperative pelvic radiographs: a comparative retrospective cohort study between DDH and non-DDH patients. Heliyon. 2024;10: e31141.
11. Karnik A, Lawande A, Lawande MA, Patkar D, Aroojis A, Bhatnagar N. Practice essentials of imaging in early diagnosis of DDH. Indian J Orthop. 2021;55:1466–79.
12. Bakarman K, Alsiddiky AM, Zamzam M, Alzain KO, Alhuzaimi FS, Rafiq Z. Developmental Dysplasia of the Hip (DDH): etiology, diagnosis, and management. Cureus. 2023;15:e43207.
13. Wu Q, Ma H, Sun J, Liu C, Fang J, Xie H, Zhang S. Application of deep-learning-based artificial intelligence in acetabular index measurement. Front Pediatr. 2023;16:10.
14. Li Q, Zhong L, Huang H, Liu H, Qin Y, Wang Y, Zhou Z, Liu H, Yang W, Qin M, Wang J, Wang Y, Zhou T, Wang D, Wang J, Xu M, Huang Y. Auxiliary diagnosis of developmental dysplasia of the hip by automated detection of Sharp's angle on standardized anteroposterior pelvic radiographs. Med (Baltim). 2019;98:e18500.
15. Wang X, Li W, Li T. Differences between Supine and Standing X-ray films in patients with DDH(Chinese). J Clin Radiol. 2022;41:2106–10.
16. Xu W, Shu L, Gong P, et al. A deep-learning aided diagnostic system in assessing developmental dysplasia of the hip on pediatric pelvic radiographs. Front Pediatr. 2022;9: 785480.
17. Zhang SC, Sun J, Liu CB, et al. Clinical application of artificial intelligence-assisted diagnosis using anteroposterior pelvic radiographs in children with developmental dysplasia of the hip. Bone Joint J. 2020;102–B:1574–81.
18. Sheridan GA, Nagle M, Howells C, et al. A radiographic clinic for developmental dysplasia of the hip (DDH). Ir J Med Sci. 2020;189:27–31.
19. Sicheng Z, Jun S, Chuanbin L, et al. Clinical application of artificial intelligence-assisted diagnosis using anteroposterior pelvic radiographs in children with developmental dysplasia of the hip. Bone Joint J. 2020;102–B:1574–81.
20. Guangyao Y, Yaoxian J, Tong L, et al. A semi-automatic diagnosis of hip dysplasia on x-ray films. Front Mol Biosci. 2020;17:613878.
21. Qiang L, Lei Z, Hongnian H, et al. Auxiliary diagnosis of developmental dysplasia of the hip by automated detection of Sharp's angle on standardized anteroposterior pelvic radiographs. Med (Baltim). 2019;98:e18500.
22. Long W, XiaoChun Y, Jiang W, et al. Short-term outcome of artificial intelligence-assisted preoperative three-dimensionalplanning of total hip arthroplasty for developmental dysplasia of the hip compared to traditional surgery. Jt Dis Relat Surg. 2023;34:571–82.
23. Fengyu L, Chihchi C, Chitung C, et al. Automatic hip detection in anteroposterior pelvic radiographs a label less practical framework. J Pers Med. 2021;11:522.
24. Wei L, Yu W, Tao J, et al. Landmarks detection with anatomical constraints for total hip arthroplasty preoperative measurements. International Conference on Medical Image Computing and Computer Assisted Intervention. 2020:670–679.
25. Harris JD, Lewis BD, Park KJ. Hip Dysplasia. Clin Sports Med. 2021;40:271–88.
26. Terje Terjesen. The natural history of acetabular dysplasia and later total hip arthroplasty in late-detected DDH: 48 patients with closed reduction followed to a mean age of 62 years. Acta Orthop. 2023;94:152–7.
27. J Zhang, C Li, M An, et al. Chinese expert consensus on the diagnosis and treatment of borderline developmental dysplasia of the hip (2022 Edition). Chin Med J (Engl). 2023;136:1012–1014.
28. Guo J, Bao W, Wang J, et al. A comprehensive evaluation framework for deep model robustness. Pattern Recogn. 2023;137: 109308.
29. M Han, G Yao, W Zhang, etal. Segmentation of CT thoracic organs by multi-resolution VB-nets. 2019. SegTHOR@ISBI.
30. Jia Z, Chunbao L, Mingyang A, et al. Chinese expert consensus on the diagnosis and treatment of borderline developmental dysplasia of the hip (2022 Edition). Chin Med J (Engl). 2023;136:1012–4.
31. Polkowski GG, Clohisy JC. Hip biomechanics. Sports Med Arthrosc Rev. 2010;18:56–62.
32. Flack NA, Nicholson HD, Woodley SJ. A review of the anatomy of the hip abductor muscles, gluteus medius, gluteus minimus, and tensor fascia lata. Clin Anat. 2012;25:697–708.
33. Scott Y, Natalie Z, Elizabeth L, et al. Developmental dysplasia of the hip. Pediatrics. 2019;143:e20181147.
34. Guoyue Yang Y, Li D, Luo C, Hui K, Xiao H, Zhang. Differences of anteroposterior pelvic radiographs between supine position and standing position in patients with developmental dysplasia of the hip. Orthop Surg. 2019;11:1142–8.

35. Yinjun L, Guangyao Y, Yaoxian J, et al. Measuring lateral center edge angle and sharp angle of hip joint on X-ray films based on deep learning for diagnosing hip dysplasia. Chin J Med Imaging Technol(Chinese). 2022;38(11):1710–4.
36. Yaoxian J, Guangyao Y, Yuan L, et al. Computer-aided system application value for assessing hip development. Front Physiol. 2020;11:587161.
37. Qiang L, Wenzhuo Y, Meng X, et al. Model construction and application for automated measurement of CE angle on pelvis orthograph based on MASK-R-CNN algorithm. Biomed Phys Eng Express. 2021;7(3):035010.
38. Holden A, Seth R, Ahmed A, et al. Artificial intelligence-generated hip radiological measurements are fast and adequate for reliable assessment of hip dysplasia: an external validation study. Bone Jt Open. 2022;3:877–84.
39. Jinyan W, Xiaodong C. Recent research progress of hip-preserving treatment for adolescents and adults with developmental dysplasia of the hip. Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi. 2021;35:1513–8 .(Chinese).
40. Hoaglund F, Kalamchi A, Poon R, et al. Congenital hip dislocation and dysplasia in Southern Chinese. Int Orthop. 1981;4:243–6.
41. KH Hsu, WC Chang, CK Feng, YP Su. Implementing the AAOS guidelines for screening of developmental dysplasia of the hip before the age of 6 months in Taiwan. J Pediatr Orthop 2023;43:e416-e420.

Publisher's Note