BMC Pregnancy and Childbirth

RESEARCH

Open Access

CLP-Net: an advanced artificial intelligence technique for localizing standard planes of cleft lip and palate by three-dimensional ultrasound in the first trimester

Guangzhi He[1,2†], Zhou Li[1,2†], Zhiyuan Zhu[3†], Tong Han[3], Yan Cao[4], Chaoyu Chen[3], Yuhao Huang[3], Haoran Dou[5,6], Lianying Liang[2], Fangmei Zhang[2], Jin Peng[3], Tao Tan[7], Hongmei Liu[8*], Xin Yang[3*] and Dong Ni[3*]

Abstract

Background Early diagnosis of cleft lip and palate (CLP) requires a multiplane examination, demanding high technical proficiency from radiologists. Therefore, this study aims to develop and validate the first artificial intelligence (AI)-based model (CLP-Net) for fully automated multi-plane localization in three-dimensional(3D) ultrasound during the first trimester.

Methods This retrospective study included 418 (394 normal, 24 CLP) 3D ultrasound from 288 pregnant woman between July 2022 to October 2024 from Shenzhen Guangming District People's Hospital during the $11–13^{+6}$ weeks of pregnancy. 320 normal volumes were used for training and validation, while 74 normal and 24 CLP volumes were used for testing. Two experienced radiologists reviewed three standard lip and palate planes (mid sagittal, retronasal triangle, and maxillary axial planes) as ground truth (GT) and the CLP-Net was developed to locate these planes.

Results In normal test set, mean angle(±SD)° and distance(±SD)mm differences were 6.24 ± 4.83, 9.81 ± 5.48, 15.36 ± 18.14 and 0.86 ± 0.72, 1.36 ± 1.15, 1.96 ± 2.35 for MSP ± SD, RTP ± SD and MAP ± SD, NCC and SSIM were 0.931 ± 0.079, 0.819 ± 0.122, 0.781 ± 0.157 and 0.896 ± 0.058, 0.785 ± 0.076, 0.726 ± 0.088 respectively. In the CLP cases, there were 8.61 ± 5.52, 10.67 ± 5.08, 16.91 ± 17.42 and 1.03 ± 1.20, 1.17 ± 1.08, 1.34 ± 0.95 for mean angle and distance in MSP, RTP, and MAP, respectively. NCC and SSIM were 0.876 ± 0.104, 0.803 ± 0.084, 0.793 ± 0.089 and 0.841 ± 0.105, 0.812 ± 0.085, 0.764 ± 0.100, respectively. CLP-Net predictions had a highly visual acceptance rate among radiologists

†Guangzhi He, Zhou Li and Zhiyuan Zhu contributed equally to this work.

*Correspondence:
Hongmei Liu
986528134@qq.com
Xin Yang
xinyang@szu.edu.cn
Dong Ni
nidong@szu.edu.cn

Full list of author information is available at the end of the article

(MSP: 95%, RTP: 70%, MAP: 70%), with improved localization speed 15s(31.3%) for senior radiologists and 63s(38.9%) for junior radiologists.

Conclusions CLP-Net accurately locates three planes for CLP screening, aiding radiologists and enhancing the efficiency of ultrasound examinations.

Keywords Artificial intelligence, Cleft lip and palate, First-trimester, 3D ultrasound, Fetus, Standard plane localization

Background

Cleft lip and palate (CLP) is a common congenital facial deformity in fetuses, with a global prevalence of 1.08 per 1,000 births [1, 2]. These deformities can cause speech, feeding, and mental and social development problems, causing distress for patients and families and resulting in a substantial societal impact [3]. In routine antenatal cares, most screenings and treatment plan for CLP are conducted in the second trimester. Extensive research [4–6] indicates that screening and diagnosis of CLP can be conducted in the first trimester of pregnancy, diagnosing CLP in the first trimester enhances treatment planning by providing more time for thorough preparation and decision-making. Furthermore, it also provides additional possibilities for women in underdeveloped areas who may face economic challenges.

Ultrasound has become the most widely used method for antenatal examination in pregnant women because it is noninvasive, convenient, and rapid. However, basic two-dimensional (2D) ultrasound scanning is limited to acquiring images in a single plane. Due to the complex developmental processes and deeper location of the palate within the facial structure, accurate localization of multiple planes is required to fully assess the condition of the fetal lip and palate [7, 8]. Furthermore, the small fetal facial structures in the first trimester, improper positioning during the scanning period, and limited knowledge of fetal lip and palate development make it challenging to obtain high-quality ultrasound images of the lip and palate in the first trimester. According to the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) [9], clinical practice in the first trimester should include an examination of facial clefts in 2D mid sagittal plane. If an abnormality is detected, axial and coronal plane investigation is necessary to further ascertain the palate structure. A related study [10] also demonstrated that diagnosing CLP requires evaluation across multiple planes using 3D ultrasound, which undoubtedly poses a challenge for junior radiologists.

In recent years, significant advancements have been made in ultrasound imaging, particularly the 3D ultrasound imaging. Previous studies have shown that 3D ultrasound can better visualize fetal facial structures and enhance the accuracy of detecting CLP [11–13]. Furthermore, compared with 2D ultrasound, 3D ultrasound can capture multiple imaging planes in a single scan, providing more comprehensive information about the fetal facial anatomy. Integrating 3D ultrasound with 2D slices for the mid sagittal plane (MSP), retronasal triangle plane (RTP), and maxillary axial plane (MAP) within the 3D ultrasound can lead to a more precise diagnosis of CLP and reduce the rate of missed diagnoses [13, 14]. Figure 1 shows the three planes and their relative positions in the 3D ultrasound. Despite the advantages, manually identifying standard planes for the lip and palate remains a time-consuming and challenging task due to the large amount of information and the orientation uncertainty associated with the 3D ultrasound.

With the development of artificial intelligence (AI) techniques, simpler and faster methods for locating standard planes using 3D ultrasound have emerged. In 2018, Chykeyuk et al. [15] used random forest to extract standard planes from 3D echocardiography. Li et al. [16] proposed a method for the autonomous navigation of an ultrasound probe towards standard lumbar vertebrae planes using reinforcement learning (RL) in 2021. Yang et al. [17–19] conducted extensive research on the localization of standard planes using 3D ultrasound of uterus and fetal brain from 2021 to 2022. Many previous studies have proven that AI technology can quickly and automatically locate multiple standard planes, reduce subjective variability, and promote the standardization and normalization of tasks involving the localization of standard planes. Nevertheless, no AI has been developed to perform the task of localizing the standard plane of fetal lip and palate, this item remains a challenging one.

Therefore, this study developed the AI-based multiplane localization model named CLP-Net to automatically locate the standard lip and palate planes in 3D ultrasound during the first trimester. This study aims to assist radiologists in quickly identifying standard planes by combining 2D and 3D information, thereby facilitating the localization of CLP and improving the efficiency of the examination in the first trimester.

Methods

Study subjects, design, and population

This retrospective study reviewed ultrasound examinations conducted at Shenzhen Guangming District People's Hospital between July 2022 and October 2024. The inclusion criteria were as follows: (1) no history of hypertension, diabetes, or other diseases may cause blurry

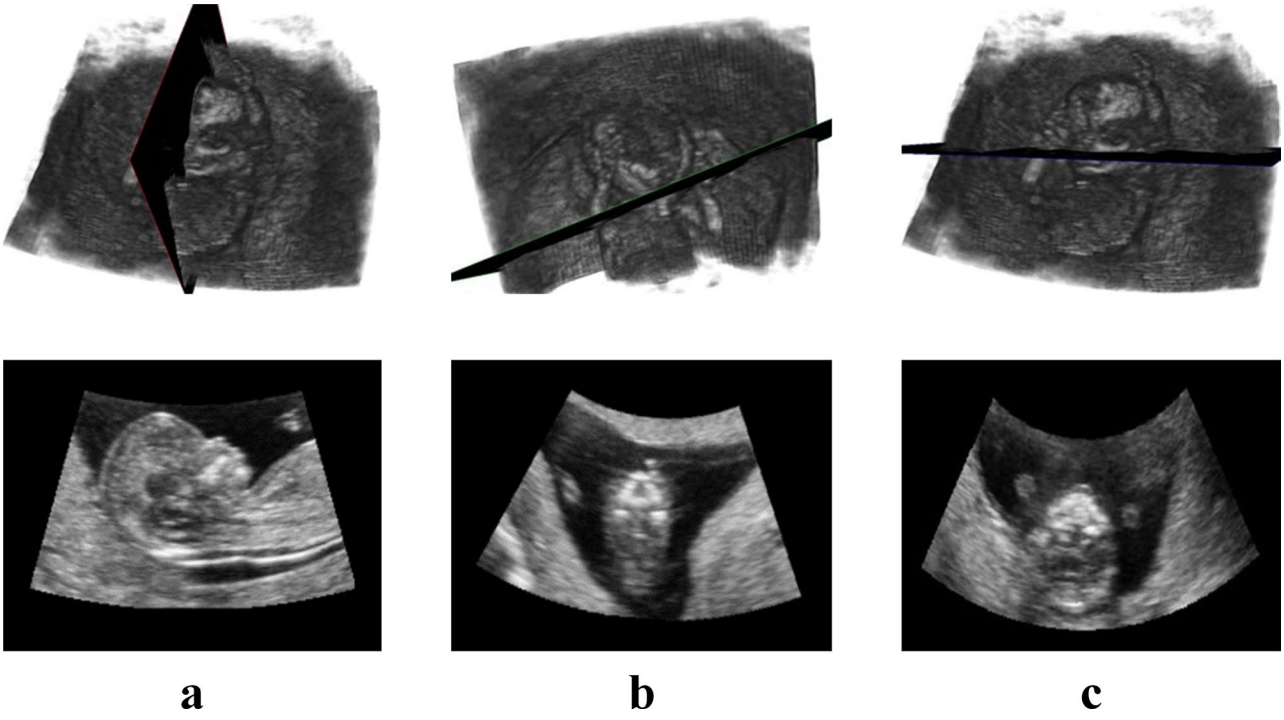

Fig. 1 illustrates, for the first time, the three target standard planes of the fetus at 11–13^{+6} weeks in 3D ultrasound. Columns (**a**), (**b**) and (**c**) depict the spatial location(upper) and 2D slice images(lower) of MSP, RTP, and MAP, respectively

volume before pregnancy, (2) singleton pregnancy, with fetal crown-rump length (CRL) of 45–84 mm and nuchal translucency (NT)≤3 mm, (3) no abnormalities were detected during the 11–13^{+6} weeks examination.

This study was approved by the Medical Ethics Committee of People's Hospital of Guangming District, Shenzhen, China (approval number: LL-KT-2023090). All pregnant women provided informed consent, and the examinations were conducted according to the ISUOG guidelines for 11–14 weeks ultrasound scan. A total of 418 high-quality 3D ultrasound from 288 pregnant women were obtained based on inclusion and annotation standards. These volumes included 24 cases of CLP obtained by 13 pregnant women and 394 normal cases from 275 pregnant women. For the normal volumes, 320 cases were selected as the training and validation sets, using four-fold cross validation to obtain the best model. The remaining 74 cases were combined with 24 CLP cases to create a dataset for comparing the spatial differences, similarities, and visual analyses between CLP-Net and ground truth (GT). Figure 2 illustrates the flowchart of the study design.

Equipment, software, and quality control

All data were obtained using the GE Voluson E10 color Doppler ultrasound diagnostic system (General Electric Company, Boston, USA). This system was equipped with a 3D volumetric probe (RM6C) operating at frequencies of 2–8 MHz. The scan measures used in the fetal assessment included estimating the CRL and NT thickness, assessing the fetal nasal bone, and conducting a comprehensive scan of major fetal structures, such as the head, face, spine, heart, thoracic and abdominal cavities, thoracoabdominal wall, kidneys, bladder, and limbs. The placenta, amniotic fluid, umbilical cord, fetal appendages, and cervix were observed.

After completing the initial assessments, the radiologist switched to a 3D volumetric transducer to acquire 3D ultrasound data. The image quality was set to 'Quality High MAX,' and the imaging frame included the fetal head and upper chest. 3D ultrasound imaging began from the mid sagittal plane when the fetus was in a natural and calm state. The expectant mother was instructed to hold her breath during the capture to minimize motion artifacts. The data were exported and saved in the uncompressed 'Voluson Format (*.4dv)' for subsequent analysis and processing. All ultrasound acquisition procedures adhere to ALARA principles.

Markers and annotations

Two senior radiologists with extensive ultrasound scanning experience annotated the volume. One radiologist has 16 years of experience and holds the Chinese Fetal Medicine Foundation (CFMF) certification for Obstetric ultrasound performed the initial annotations, while the other has 30 years of experience and also possesses CFMF certification reviewed and verified the results. When discrepancies arose, two radiologists discussed

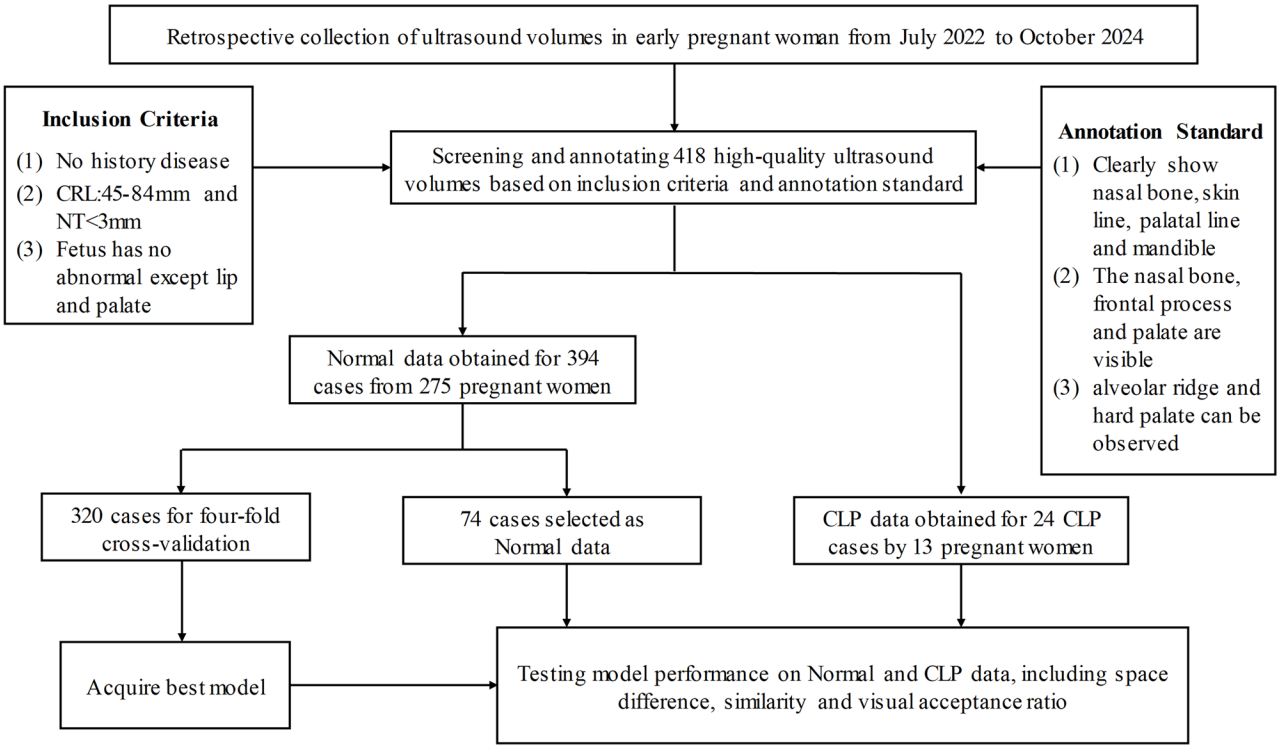

Fig. 2 Flowchart summarizing the study design

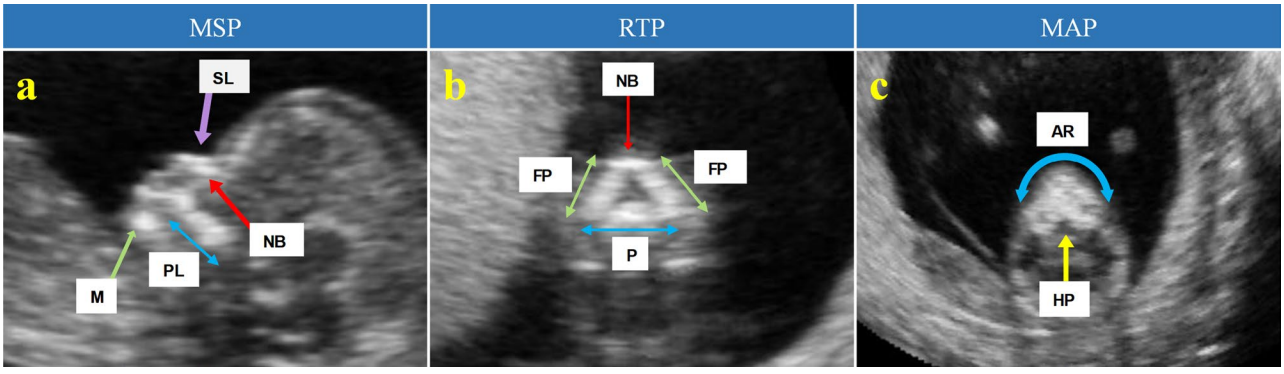

Fig. 3 Illustrates the structural features of the standard planes. (**a**) Skin line (SL), nasal bone (NB), palatal line (PL), and mandible (M) features of the MSP. (**b**) Nasal bone (NB), frontal process (FP), and palate (P) of the RTP. (**c**) Alveolar ridge (AR) and hard palate (HP) on MAP

and revised the annotations together until a consensus was reached. The "Pair" [20] software (version: 2.6.0), independently developed by Shenzhen RayShape Medical Technology Co., Ltd (Shenzhen, Guangdong, China) was used to annotate the exported 3D ultrasound data. The following planes were annotated in this study:

1) MSP: The facial structure in a sagittal view, including the nasal bone, surface skin line, palate line, and mandible.
2) RTP: Displays the nasal bone, frontal process, and upper alveolar process clearly in a coronal view.
3) MAP: Reveals smooth and intact structures of the upper alveolar ridge and hard palate in an axial view.

Figure 3 shows the structures of the three planes mentioned above

Development of the CLP-Net

The framework of the proposed CLP-Net based on RL is illustrated in Fig. 4. This work extends the findings of Zou et al. [17], where Convolutional Neural Networks (CNNs) were used as Agents, and spatial-anatomical similarity as Reward. The Agents interact with the volume environment, taking actions to predict plane coordinates based

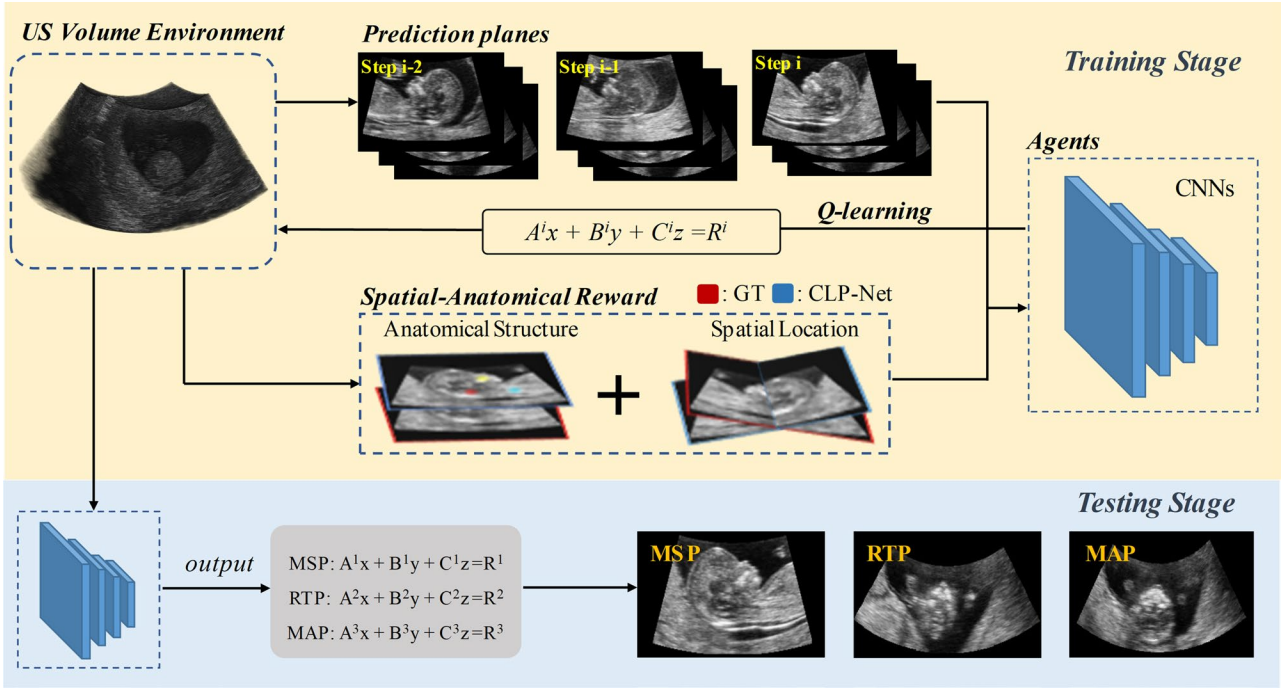

Fig. 4 Architecture of RL-based CLP-Net for standard plane localization in 3D ultrasound

on the Q-learning algorithm. We applied the proposed method to the task of locating the standard planes of the lip and palate. In the training stage, CLP-Net moved the prediction planes to approach the GTs iteratively. This process involves the following steps: (1) Agents took actions based on the image features from the previous three steps, to iteratively interact with the environment to output plane coordinates and reconstructed slices at each step; (2) The environment will give feedback to the agents via anatomical structure and spatial location rewards calculated between the GT and prediction planes; (3) The agent-environment interaction will stop based on the pre-set termination strategy (i.e., reaching 60 maximum steps). In the testing stage, the agents can directly exploit the knowledge learned during the training to determine the coordinates of the multiple standard planes of MSP, RTP and MAP.

Statistical analysis

Statistical analysis was performed using SPSS (version 25.0; Chicago, Illinois, USA) to assess the spatial differences between CLP-Net and GT. This evaluation involved the Euclidean distance between center point coordinates of the GT and CLP-Net predictions, as well as the angle between two planes within the 3D ultrasound. The study also compared the similarity using two metrics: normalized cross-correlation [21] (NCC) and the structural similarity index measure [22] (SSIM). NCC is a method that utilizes grayscale information to assess the similarity between two images which can be expressed as:

$$\mathrm{NCC}(X,Y) = \frac{\sum (X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum (X-\bar{X})^2 \sum (Y-\bar{Y})^2}}, \quad (1)$$

Where X and Y represent the two images used to calculate similarity, $\bar{X}$ and $\bar{Y}$ represent their respective grayscale means. SSIM is a metric designed to quantify image similarity by examining luminance, contrast, and structural information. Compared to traditional metrics, SSIM aligns more closely with human perceptual patterns. SSIM is defined as:

$$SSIM(X,Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)}, \quad (2)$$

Where X and Y represent the two images used to calculate similarity, μ_X, μ_Y and σ_Y, σ_Y represent the mean and variance of X and Y respectively, σ_{XY} is the covariance of X and Y. C_1, C_2 are constants to avoid the denominator being zeros. Additionally, to evaluate the acceptability of CLP-Net, a radiologist with extensive clinical experience assessed the visual acceptance ratio of the CLP-Net prediction planes. For all analyses, the chi-square test was used for categorical variables, and the t-test was used for continuous variables. Statistical significance was set at $p<0.05$.

Table 1 Demographic characteristics of pregnant women and fetus

Variable	Normal data ($n=275$)	CLP data ($n=13$)	*P*-value
Pregnant Women			
Age (years)	29.3 ± 4.7	27.9 ± 5.7	0.167
Cesarean			0.428
Yes	34(12.4%)	2(18.2%)	
No	241(87.6%)	11(91.8%)	
Gravida			0.183
1	105(38.2%)	9(69.2%)	
2	119(43.3%)	1(7.7%)	
≥ 3	51(19.5%)	3(23.1%)	
Placenta location			0.833
Anterior	128(46.6%)	7(53.8%)	
Posterior	142(51.6%)	6(46.2%)	
fundal	5(1.8%)	0(0%)	
Fetus			
GA (days)	90.1 ± 2.8	90.23 ± 2.86	0.524
NT (mm)	1.6 ± 0.4	1.87 ± 0.72	0.486
CRL (mm)	65.4 ± 5.4	64.23 ± 8.52	0.325

*GA Gestational Age,

*NT Nuchal Translucency, *CRL Crown-Rump Length

Results

A total of 288 women aged 18–47 years in their first-trimester ($11–13^{+6}$ weeks) were included in this study. The demographic characteristics of the participants are shown in Table 1. P-value analysis showed that the data were not significantly different.

We constructed our model using PyTorch 2.1.0 and trained it for 100 epochs (spend ten hours) on an NVIDIA RTX 2080Ti GPU. The optimizer used was Adam with a learning rate of 1e-4, which linearly decayed by 10% every 30 epochs. Additionally, the 3D ultrasound data is reshaped into a numpy array format of (1, 224, 224, 224) before training, where the first dimension represents the grayscale channel and the remaining dimensions correspond to the image's width, height, and depth. The grayscale values are normalized to a range between 0 and 1.

Performance of CLP-Net on the normal data

Table 2 presents the comparative results of differences and similarities between the two groups.

To obtain more stable results from the model, all result were the average of the predicted results from the test set based on the model obtained through four-fold cross-validation. Specifically, the mean angle differences of MSP ± SD, RTP ± SD, and MAP ± SD were 6.24 ± 4.83°, 9.81 ± 5.48°, and 15.36 ± 18.14°, respectively, while the corresponding distances were 0.86 ± 0.72 mm, 1.36 ± 1.15 mm, and 1.96 ± 2.35 mm. To comprehensively assess localization accuracy, we conducted similarity analyses using NCC and SSIM to evaluate the consistency of anatomical structures between CLP-Net and GT. The NCC values obtained were 0.931 ± 0.079 for MSP, 0.819 ± 0.122 for RTP, and 0.781 ± 0.157 for MAP. The corresponding SSIM values were 0.896 ± 0.058, 0.785 ± 0.076, and 0.726 ± 0.088, respectively. To demonstrate the superiority of our method, we conducted a more comprehensive analysis of the model, including a comparison with the 3D-ResNet50 baseline method. Additionally, to better illustrate the impact of each component of the model on the results, we conducted an ablation study on the model's anatomical structure(AS) and spatial location(SL) reward components. The results of these experiments and visualization are presented in Table 2, Fig. 5.

Performance of CLP-Net on CLP data

To assess the localization accuracy of CLP-Net on CLP samples, we calculated the space differences and similarities between CLP-Net and GT on the CLP data. Table 3 provides details of the comparison. MSP exhibited superior performance in both angle and distance differences,

Table 2 Comparison analysis of the CLP-Net and ablation study

		3D-Resnet50 (baseline)	CLP-Net (w/o AS)	CLP-Net (w/o SL)	CLP-Net
MSP	Angle(°) ↓	28.72 ± 6.26	10.23 ± 5.32	6.93 ± 4.11	**6.24** ± 4.83
	Distance (mm) ↓	0.84 ± 0.83	0.91 ± 0.35	**0.78** ± 0.67	0.86 ± 0.72
	NCC ↑	0.717 ± 0.051	0.824 ± 0.039	0.897 ± 0.069	**0.931** ± 0.079
	SSIM ↑	0.786 ± 0.045	0.801 ± 0.063	0.877 ± 0.061	**0.896** ± 0.058
RTP	Angle(°) ↓	54.70 ± 16.486	15.37 ± 4.36	11.67 ± 6.29	**9.81** ± 5.48
	Distance (mm)↓	4.37 ± 2.69	2.13 ± 1.06	1.57 ± 1.03	**1.36** ± 1.15
	NCC ↑	0.598 ± 0.082	0.768 ± 0.094	0.807 ± 0.820	**0.819** ± 0.122
	SSIM ↑	0.673 ± 0.074	0.742 ± 0.015	0.784 ± 0.055	**0.785** ± 0.076
MAP	Angle(°) ↓	51.63 ± 17.74	19.28 ± 3.31	16.2 ± 14.9	**15.36** ± 18.14
	Distance (mm) ↓	2.81 ± 2.40	**1.56** ± 1.13	1.82 ± 1.71	1.96 ± 2.35
	NCC ↑	0.667 ± 0.081	0.714 ± 0.036	0.750 ± 0.130	**0.781** ± 0.157
	SSIM ↑	0.664 ± 0.078	0.635 ± 0.049	**0.763** ± 0.076	0.726 ± 0.088

**AS*: anatomical structure **SL*: spatial location

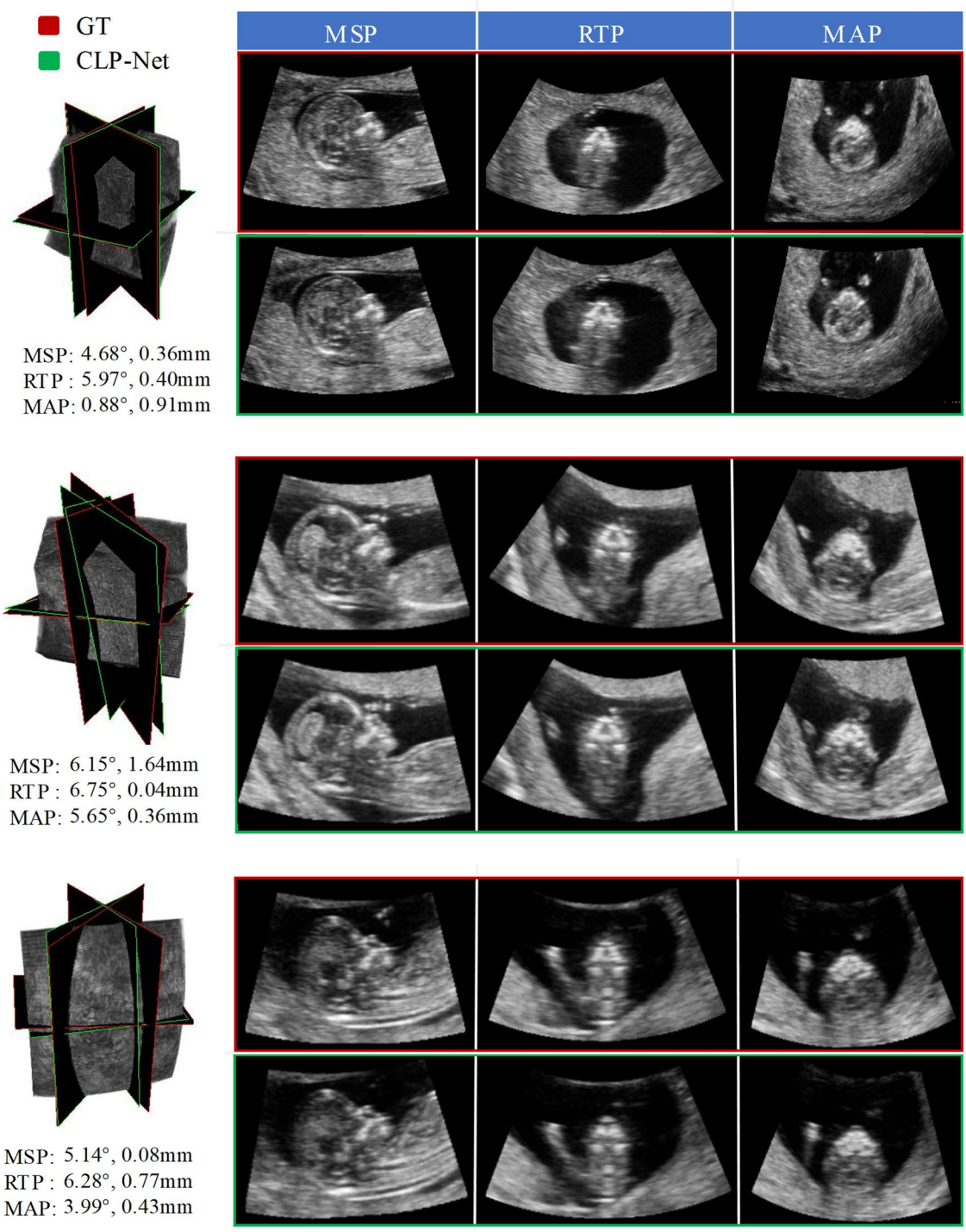

Fig. 5 Comparison of GT (red) and CLP-Net predictions (green) on the normal data. The left side of each sample shows the relative positions and differences, while the right side's three columns compare GT and AI predictions for MSP, RTP, and MAP

with values of 8.61±5.52° and 1.03±1.02 mm, respectively. For RTP and MAP, the angle differences were 10.67±5.08°, 16.91±17.42°, while the distance differences were 1.17±1.08 mm and 1.34±0.95 mm, respectively. Figure 6 illustrates the contrast between the CLP-Net prediction planes and the GT. We also evaluated the performance of CLP-Net using NCC and SSIM. The NCC values for MSP, RTP, and MAP were 0.876±0.104, 0.803±0.084, and 0.793±0.089, respectively. Correspondingly, the SSIM values were 0.841±0.105, 0.812±0.085, and 0.764±0.100, respectively. Based on these results, CLP-Net demonstrated good consistency between the

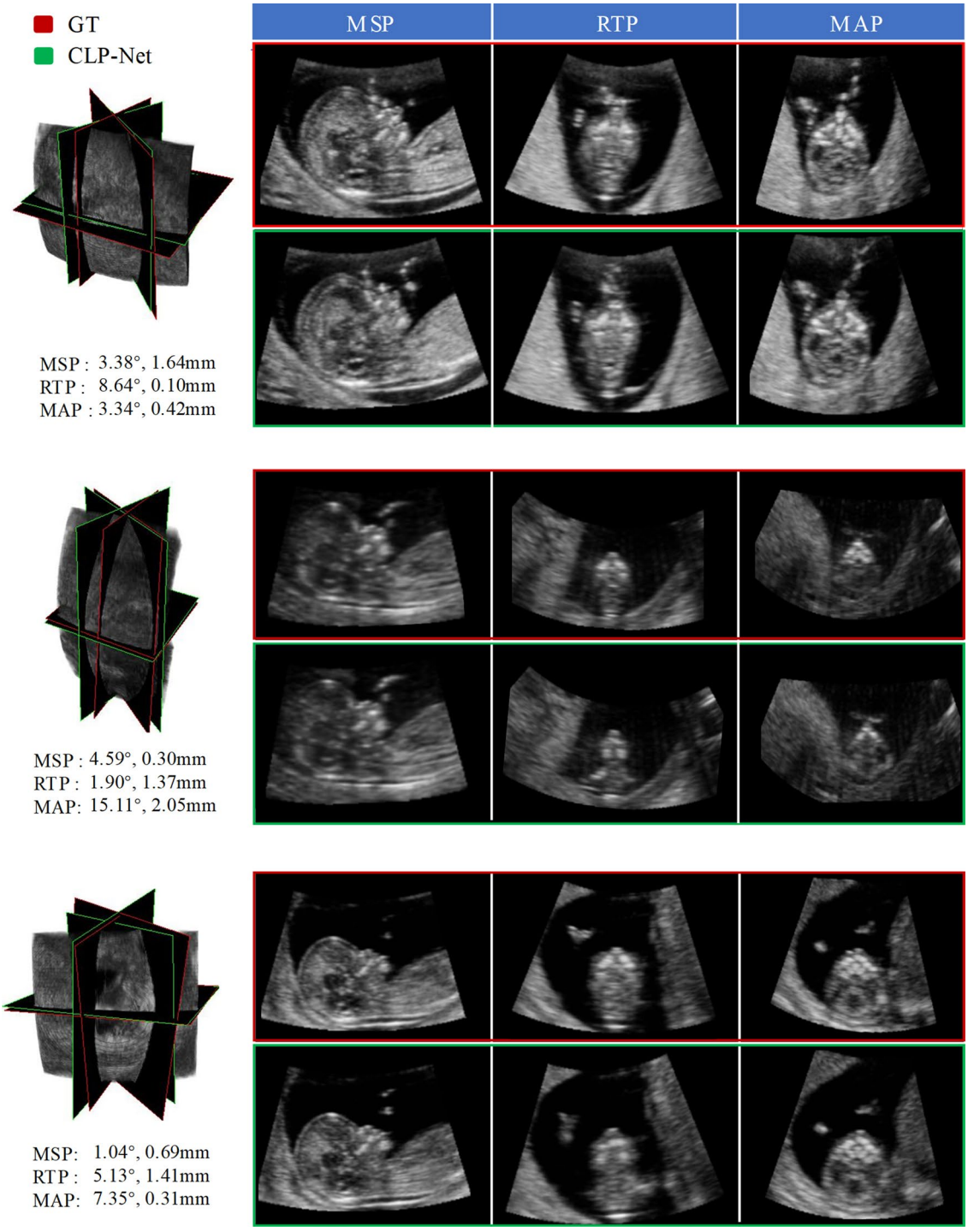

Fig. 6 The comparison of GT and CLP-Net predictions on the CLP data

Table 3 Comparison analysis of the space difference and similarity between CLP-Net and GT on CLP data

	MSP	RTP	MAP
Angle(°) ↓	8.61 ± 5.52	10.67 ± 5.08	16.91 ± 17.42
Distance(mm) ↓	1.03 ± 1.20	1.17 ± 1.08	1.34 ± 0.95
NCC ↑	0.876 ± 0.104	0.803 ± 0.084	0.793 ± 0.089
SSIM ↑	0.841 ± 0.105	0.812 ± 0.085	0.764 ± 0.100

normal and CLP data, and MSP had significant performance across the three planes.

Visual acceptability and time-analysis test

We also statistically analyzed the visual acceptance ratio of CLP-Net, focusing on visual plausibility of the localized planes. Forty 3D ultrasound were randomly selected

Table 4 Visual acceptance of CLP-Net predictions

	Randomly sample 3D ultrasound from normal data ($n=40$)		
	MSP	**RTP**	**MAP**
Acceptance	38(95%)	28(70%)	28(70%)
Rejection	2(5%)	12(30%)	12(30%)
NB-L	2(5%)	3(7.5%)	-
PL-L	0(0%)	5(12.5%)	-
Md-L	0(0%)	-	-
FP-L	-	10(25%)	-
AR-L	-	-	11(27.5%)
HP-L	-	-	9(22.5%)

**NB-L*: Nasal Bone-Loss, **PL-L*: Palatal Line-Loss,
**Md-L*: Mandible-Loss, **FP-L*: Frontal Process-Loss,
**AR-L*: Alveolar Ridge-Loss, **HP-L*: Hard Palate-Loss

Table 5 Comparison of time consumption between manual and CLP-Net methods

	Manual Location ($n=40$)	**CLP-Net Assisted ($n=40$)**	**Improvement**
Senior			
Average Time	48s	33s	15s (31.3%)
Junior			
Average Time	162s	99s	63s (38.9%)
Acceptance	84 (70.0%)	112 (93.3%)	28 (23.3%)

from a normal data for visual acceptability testing by radiologists. Table 4 displays the evaluation results. The acceptance rates for MSP, RTP, and MAP were 95%, 70%, and 70%, respectively. For cases deemed unacceptable, we conducted a retrospective review to explore the reasons for failure according to established annotation standards. Our findings indicated that MSP had high acceptance, with only two samples rejected due to nasal bone loss (NB-L). RTP and MAP had similar rejection rates. RTP was rejected more often due to frontal process loss (FP-L) (10/12), while MAP was rejected more due to alveolar ridge loss (AR-L) (11/12) and hard palate loss (HP-L) (9/12).

We further randomly selected 40 3D ultrasound, comprising 120 planes, to conduct a time-analysis test. Table 5 illustrates the average time spent by senior and junior radiologists to locate a single plane, with a prediction time of approximately 0.49 s per case. Without assistance, a senior radiologist could locate three standard planes in 3D ultrasound within 48 s, whereas a junior radiologist took 162 s and often struggled to locate standard planes. Assisted by CLP-Net, the senior radiologist could locate a volume in approximately 33 s, whereas the junior could do so in 99 s, and the visual acceptance ratio improved from 70 to 93.3%.

Discussion

CLP is the most common congenital malformation of the oral and maxillofacial regions, with a high incidence and potential for significant harm. In traditional ultrasound scanning, multisection 2D ultrasound combined with 3D ultrasound in mid-pregnancy is considered the most reliable and accurate method for examining CLP. However, we think this screening method may suffer from several drawbacks. First, CLP screening is mostly carried out in the second trimester, while fetal facial begins in the first trimester. Second, multiplane scanning requires more time, which slows down examination efficiency. Additionally, radiologists may be unfamiliar with the complex fetal facial structures, further extending scanning time.

In this study, to the best of our knowledge, we developed the first novel AI model called CLP-Net to automatically localize the standard planes of the lip and palate in 3D ultrasound during the first trimester. This model can be an effective solution to the problems associated with CLP examination. We conducted a comprehensive analysis of CLP-Net, evaluating its performance across various aspects, including spatial differences, similarities, visual acceptance ratio, localization speeds, and assistance to radiologists. CLP-Net demonstrated balanced performance across these different metrics, demonstrating its stability.

In the analysis of distance and angle differences, CLP-Net demonstrated better overall performance than other similar studies [14–16]. When comparing the prediction planes within the task, MSP showed the best performance, whereas MAP was relatively weaker. This may be due to the wider depth of the palate in the sagittal view, which allows for a larger visible interval. Additionally, acquiring the 3D ultrasound requires the fetus to be in a fixed position, which contributes to the superior performance of MSP. In contrast, the MAP and RTP may have large differences in spatial position between individuals due to movements of the fetal head, increasing the difficulty of model learning. The results of NCC and SSIM were directly proportional to angle and distance differences, which supports our analysis.

To further explore the actual performance of the model, we analyzed the visual acceptance ratio of CLP-Net, and the results were 95%, 70%, and 70% for MSP, RTP, and MAP, respectively, which means that doctors can directly diagnose most planes without time-consuming localization work.

For the mis-localized RTP and MAP planes, we analyzed the underlying causes of failure. It was observed that, in certain cases, predicted planes with minimal angular and distance errors were still deemed "unacceptable" by sonography. This discrepancy arises because clinical evaluations are based on image features rather than coordinate errors. To better meet clinical

requirements, it is essential to incorporate more semantic constraints into future AI model development, rather than spatial coordinate constraints.

In cases where there were inaccuracies in the predicted locations, radiologists could locate the correct planes faster based on the prediction results, aligning with the purpose of our research. In the final localization speed test, we found that the improvement for junior radiologists was three times that for senior radiologists, suggesting that our CLP-Net is more beneficial for junior radiologists than for senior radiologists. Additionally, the visual acceptance ratio for juniors increased from 70 to 93%. These results demonstrate that CLP-Net serves as a valuable guide for inexperienced practitioners. CLP-Net has shown great potential and application prospects for prenatal ultrasound examinations, especially in settings where experienced personnel are scarce.

However, this study has several limitations. Due to the emission of ultrasound waves by the 3D ultrasound probe's array being perpendicular to the fetal coronal plane, the MSP is parallel to the ultrasound waves, allowing for direct image acquisition. In contrast, RTP and MAP are at an angle to the ultrasound array, resulting in low-resolution images reconstructed from multiple vertical beams. This leads to suboptimal performance of the model on the RTP and MAP. Moreover, the variations in fetal facial tilt angles, both upward and downward, further increase the range of changes in RTP and MAP, further increasing the predictive difficulty. Another significant challenge was the rigorous data selection and annotation process, which can be time-consuming and resource-intensive, resulting in a limited amount of 3D ultrasound data available for model training. Additionally, due to the limited number of CLP samples, the results of the statistical analysis of the CLP samples may be biased. To obtain more robust results, additional CLP samples should be collected.

In this work, we localized the standard planes of lip and palate from 3D ultrasound and achieved favorable results. The automatic diagnosis of CLP based on standard planes represents a promising option for the future. This approach could establish a comprehensive diagnostic process, seamlessly integrating localization and diagnosis, thereby enabling underserved areas and regions with a shortage of professionals to access essential healthcare services.

Conclusions

In this study, we developed CLP-Net to automatically identify standard lip and palate planes from 3D ultrasound in the first trimester. The model's accuracy was impressive, demonstrating its potential to improve the efficiency and accuracy of ultrasound diagnoses. This could facilitate ultrasound examinations for the general public, contributing to advancements in reproductive health.

Abbreviations

CLP	Cleft Lip and Palate
AI	Artificial Intelligence
US	Ultrasound
3D	Three-dimensional
GT	Ground Truth
MSP	Mid Sagittal Plane
RTP	Retronasal Triangle Plane
MAP	Maxillary Axial Plane
NCC	Normalized Cross-correlation
SSIM	Structural Similarity Index Measure
2D	Two-dimensional
ISUOG	International Society of Ultrasound in Obstetrics and Gynecology
FMF	Fetal Medicine Foundation
CRL	Crown-rump Length
NT	Nuchal Translucency
AS	Anatomical Structure
SL	Spatial Location
SL	Skin Line
NB	Nasal Bone
PL	Palatal Line
FP	Frontal Process
AR	Alveolar Ridge
HP	Hard Palate
CNNs	Convolutional Neural Networks

Acknowledgements

Not applicable.

Author contributions

X.Y., G.H., Y.C.: Conceptualization; G.H., Z.L., Z.Z., T.H., L.L., F.Z., J.P., C.C.: Data collection and analysis; Z.L., Z.Z., Y.C.: Manuscript writing; Y.H., H.D.: Method construction; G.H., X.Y., D.N., T.T.: Funding acquisition, project administration; H.L., X.Y., and D.N.: supervision; Z.Z., Y.C., T.H., Y.H., and X.Y.: proofreading the draft. All the authors of this paper have read and approved the final submitted version.

Funding

This work was supported by grants from the National Natural Science Foundation of China (Nos. 62101343 and 62171290), Shenzhen-Hong Kong Joint Research Program (No. SGDX20201103095613036), Science and Technology Planning Project of Guangdong Province (2023A0505020002), Guangdong Yiyang Healthcare Charity Foundation (2023CSM002), Shenzhen Guangming District Health System Research Project (gmws2022019) and Science and Technology Development Fund of Macao (0021/2022/AGJ).

Data availability

The datasets and codes are not publicly available because of hospital policy and personal privacy but are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the Medical Ethics Committee of the People's Hospital of Guangming District, Shenzhen (LL-KT-2023090) and followed the tenets of the Declaration of Helsinki.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Jinan University, Guangzhou, Guangdong, China

[2]Department of Ultrasound, Shenzhen Guangming District People's Hospital, Songbai Road, Matian Street, Shenzhen, Guangdong, China
[3]National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, School of Biomedical Engineering, Health Science Center, Shenzhen University, Xueyuan Blvd, Nanshan, Shenzhen, Guangdong, China
[4]Shenzhen RayShape Medical Technology Co., Ltd, Shenzhen, Guangdong, China
[5]Department of Computer Science, School of Engineering, University of Manchester, Manchester, UK
[6]School of Computing, University of Leeds, Leeds, UK
[7]Faculty of Applied Sciences, Macao Polytechnic University, Macao, SAR, China
[8]Department of Ultrasound, Institute of Ultrasound in Musculoskeletal Sports Medicine, The Affiliated Guangdong Second Provincial General Hospital of Jinan University, Guangzhou, Guangdong, China

Received: 12 May 2024 / Accepted: 23 December 2024
Published online: 07 January 2025

References

1. Salari N, Darvishi N, Heydari M, et al. Global prevalence of cleft palate, cleft lip and cleft palate and lip: a comprehensive systematic review and meta-analysis. J Stomatology oral Maxillofacial Surg. 2022;123(2):110–20.
2. Fan D, Wu S, Liu L, et al. Prevalence of non-syndromic orofacial clefts: based on 15,094,978 Chinese perinatal infants. Oncotarget. 2018;9(17):13981.
3. Wang M, Meng R, Wang Z, et al. Prevalence of oral clefts among live births in Gansu Province, China. Int J Environ Res Public Health. 2018;15(2):380.
4. Zheng C, Ji C, Yin L et al. Ultrasonographic diagnosis of fetal cleft lip and palate during first-trimester (11–13 + 6 gestational weeks). Chin J Ultrasonography, 2021: 697–702.
5. Salazar Trujillo A, Rincón-Guio C, Lopez Narvaez L, et al. First trimester sonographic diagnosis of orofacial defects. Review of literature. J Maternal-Fetal Neonatal Med. 2020;33(18):3200–6.
6. Liao Y, Wen H, Ouyang S, et al. Routine first-trimester ultrasound screening using a standardized anatomical protocol. Am J Obstet Gynecol. 2021;224(4):396. e1-396. e15.
7. Sepulveda W, Wong AE, Martinez-Ten P, et al. Retronasal triangle: a sonographic landmark for the screening of cleft palate in the first trimester. Ultrasound Obstet Gynecol. 2010;35(1):7–13.
8. Lakshmy SR, Deepa S, Rose N, et al. First-trimester sonographic evaluation of palatine clefts: a novel diagnostic approach. J Ultrasound Med. 2017;36(7):1397–414.
9. Salomon LJ, Alfirevic Z, Berghella V, et al. ISUOG Practice guidelines (updated): performance of the routine mid-trimester fetal ultrasound scan. Ultrasound Obstet Gynecol. 2022;59(6):840–56.
10. Chaoui R, Orosz G, Heling KS, et al. Maxillary gap at 11–13 weeks' gestation: marker of cleft lip and palate. Ultrasound Obstet Gynecol. 2015;46(6):665–9.
11. Shao X, Liang L, Liu Y, et al. Comparison of diagnostic values between 2D three-section ultrasound and 3D tomographic ultrasound imaging for fetal cleft palate at $11–13^{+6}$ weeks. J Radiation Res Appl Sci. 2024;17(1):100808.
12. Ji C, Yang Z, Yin L, et al. The application of three-dimensional ultrasound with reformatting technique in the diagnosis of fetal cleft lip/palate. J Clin Ultrasound. 2021;49(4):307–14.
13. Martinez-Ten P, Adiego B, Illescas T, et al. First-trimester diagnosis of cleft lip and palate using three-dimensional ultrasound. Ultrasound Obstet Gynecol. 2012;40(1):40–6.
14. Bäumler M, Faure JM, Bigorre M, et al. Accuracy of prenatal three-dimensional ultrasound in the diagnosis of cleft hard palate when cleft lip is present. Ultrasound Obstet Gynecol. 2011;38(4):440–4.
15. Chykeyuk K, Yaqub M, Alison Noble J. Class-specific regression random forest for accurate extraction of standard planes from 3D echocardiography. Int MICCAI Workshop Med Comput Vis Springer Int Publishing. 2014: vol. 8331: 53–62.
16. Li K, Wang J, Xu Y et al. Autonomous navigation of an ultrasound probe towards standard scan planes with deep reinforcement learning. *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021: 8302–8308.
17. Zou Y, Dou H, Huang Y et al. Agent with Tangent-Based Formulation and Anatomical Perception for Standard Plane Localization in 3D Ultrasound. *International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland*. 2022: 300–309.
18. Yang X, Dou H, Huang R, et al. Agent with warm start and adaptive dynamic termination for plane localization in 3D ultrasound. IEEE Trans Med Imaging. 2021;40(7):1950–61.
19. Yang X, Huang Y, Huang R, et al. Searching collaborative agents for multi-plane localization in 3D ultrasound. Med Image Anal. 2021;72:102119.
20. Liang J, Yang X, Huang Y, et al. Sketch guided and progressive growing GAN for realistic and editable ultrasound image synthesis. Med Image Anal. 2022;79:102461.
21. Yoo JC, Han TH. Fast normalized cross-correlation. Circuits Syst Signal Process. 2009;28:819–43.
22. Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13(4):600–12.

Publisher's note