Online Resource 1 — Supplementary figures and captions for "*Cross-center sex estimation from patellar CT: multi-feature transformer fusion of morphometry, radiomics, and deep features*"

Author Names and Affiliations: Tao Liu[1], Xiurong Li[2], Haihua Bao[3], Zhanjin Wang[4], Zhiyao Yu[1], Zhan Wang[5],*, Haiyan Wang[6],*

[1] Department of Basic Medical Sciences, Qinghai University Medical College, No. 251 Ningda Road, Xining 810016, Qinghai, China

[2] Qinghai Renji Hospital, No. 172 Xincheng Avenue, Chengzhong District, Xining 810016, Qinghai, China

[3] Department of Medical Imaging Center, Qinghai University Affiliated Hospital, No. 29 Tongren Road, Xining 810001, Qinghai, China

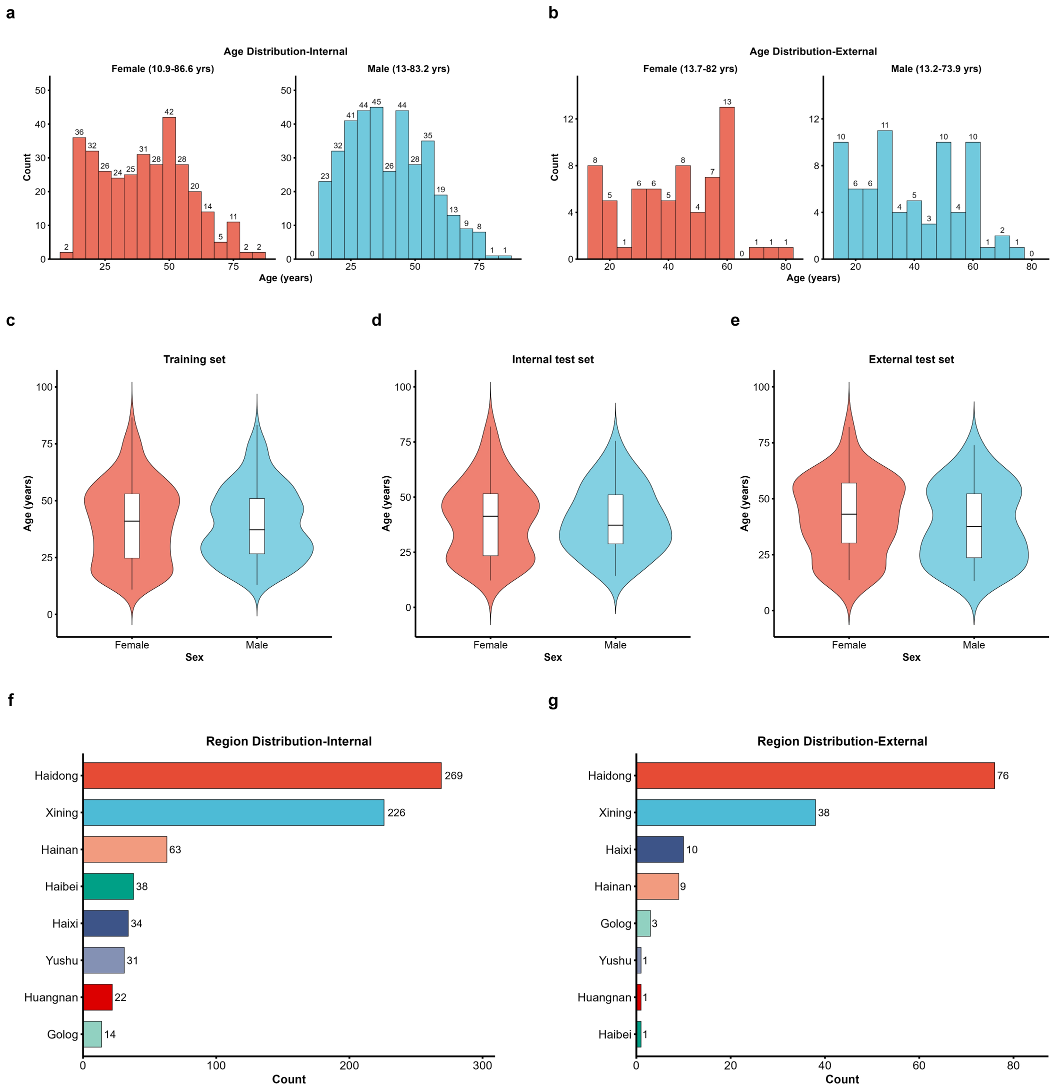
[4] Qinghai University, No. 251 Ningda Road, Xining 810016, Qinghai, China

[5] Department of Hepatobiliary and Pancreatic Surgery, Qinghai University Affiliated Hospital, No. 29 Tongren Road, Xining, Qinghai, China; Department of Medical Engineering and Translational Applications, Qinghai University Affiliated Hospital, Xining 810016, Qinghai, China

[6] Joint Research Center of Engineering Medicine for High-Altitude Cardio Cerebrovascular Health, No. 251 Ningda Road, Xining, Qinghai, China; Department of Basic Medical Sciences, Qinghai University Medical College, Xining 810016, Qinghai, China

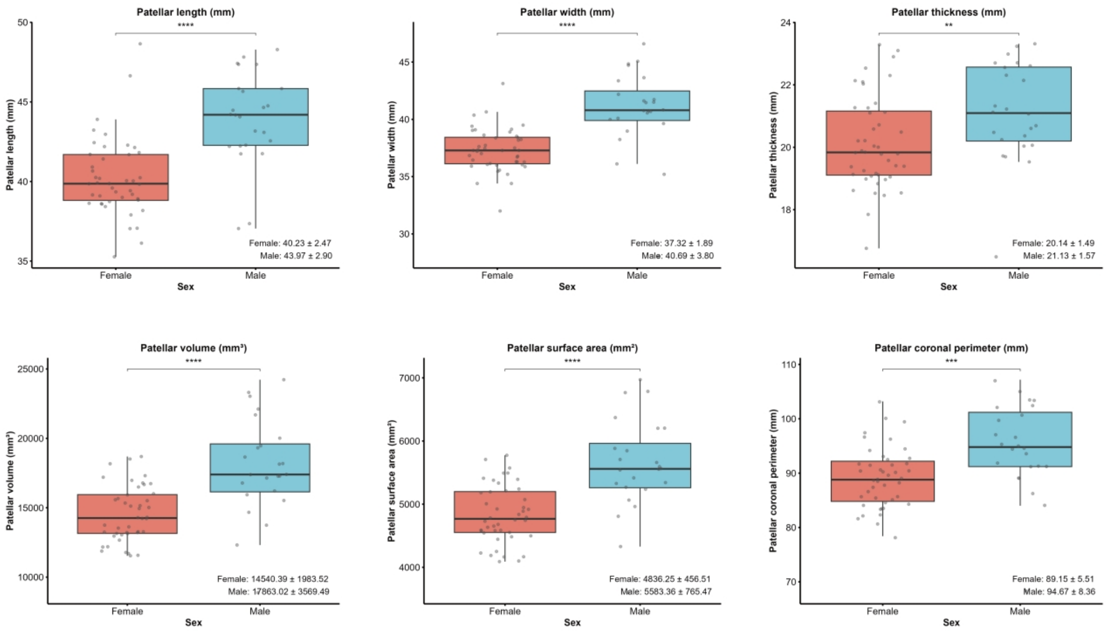
* Corresponding author.

Fig. S1 Demographic overview of the patella CT dataset.



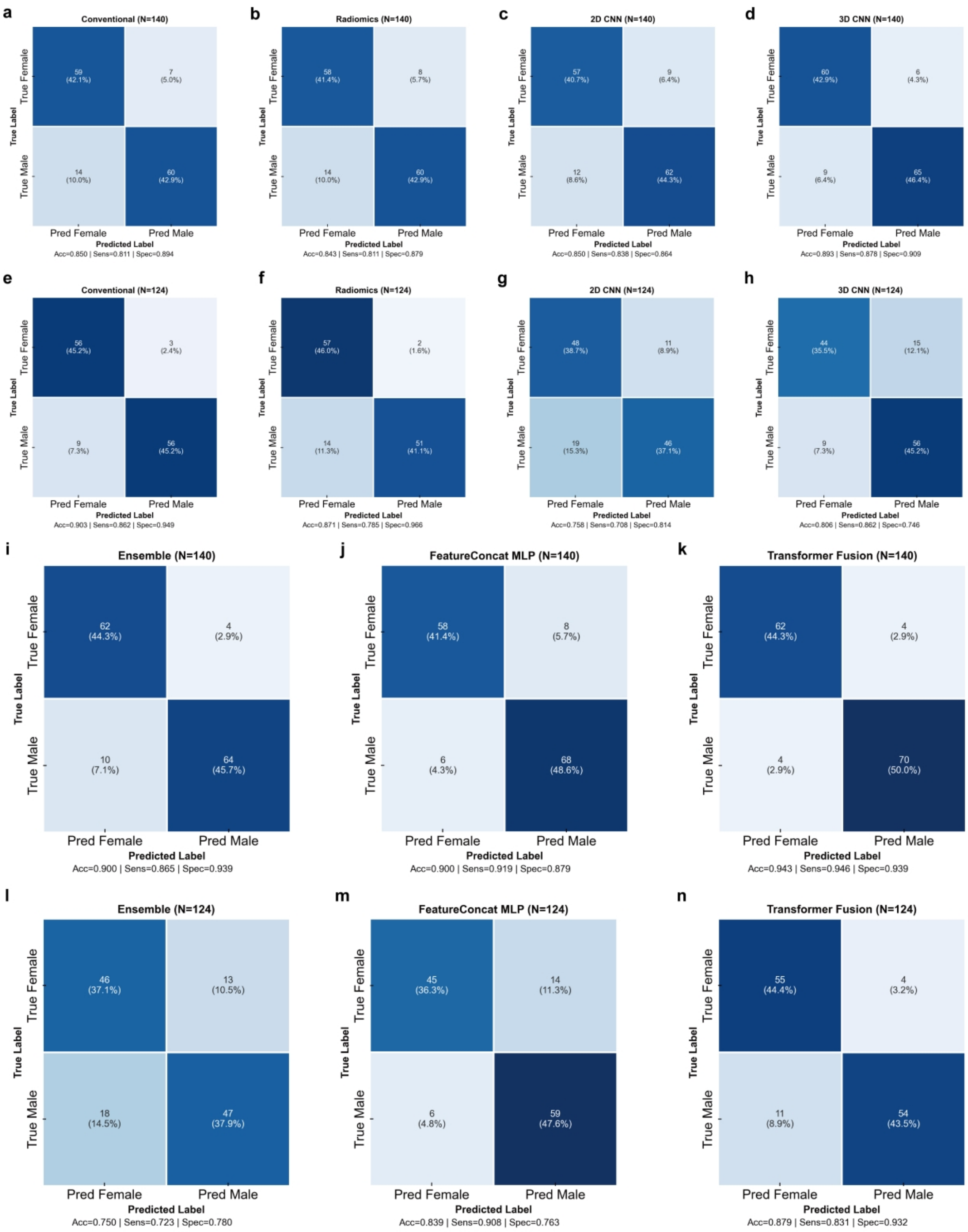
a Age distribution of the internal cohort stratified by sex. **b** Age distribution of the external cohort stratified by sex. **c–e** Violin plots of age with overlaid boxplots (median, interquartile range). c Training set. d Internal test set. e External cohort. **f**, **g** Geographic distribution of samples by prefecture-level division in the study region. **f** Internal cohort. **g** External cohort.

Fig. S2 Sex-stratified comparison of six patellar morphometric parameters in the internal cohort for individuals under 18 years of age.



Boxplots with overlaid jittered points depict the distribution of patellar length, width , thickness, volume, surface area, and coronal perimeter between female (n = 44) and male (n = 22) participants. All measurements were automatically derived from three-dimensional CT reconstructions. The mean ± standard deviation for each sex is annotated at the lower right of each panel. Statistical significance was assessed using two-sided Wilcoxon rank-sum tests; significance levels are indicated as $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, ns, not significant. Samples were drawn exclusively from the internal dataset after filtering for age < 18 years.

Fig. S3 Confusion matrices of all models.



a–d Single-feature models on the internal test set. **a** Conventional. **b** Radiomics. **c** 2D CNN. **d** 3D CNN. **e–h** Single-feature models on the external test set. **e** Conventional. **f** Radiomics. **g** 2D CNN. **h** 3D CNN. **i–k** Fusion models on the internal test set (n = 140). **i** Ensemble. **j** FeatureConcat MLP. **k** Transformer fusion. l–n Fusion models on the external test set. **l** Ensemble. **m** FeatureConcat MLP. **n** Transformer fusion. Each panel reports the number of true and predicted cases; accuracy, sensitivity and specificity are annotated below each matrix.

Supplementary Tables

Supplementary Table 1 Individual-level baseline characteristics of the training, internal test and external validation cohorts

Each row corresponds to one study participant. The full dataset is provided in Online Resource 2 (sheet Supplementary Table 1). The cohort is divided into three subsets: training set (n = 557), internal test set (n = 140), and external validation set (n = 139). Variables for each participant include unique ID, sex (1 = male, 0 = female), age at CT scanning (years), and knee laterality (1 = left, 2 = right).

Supplementary Table 2 Detailed automated patellar morphometric parameters for all study cases in the internal and external cohorts.

Each row corresponds to one study case. The full dataset is provided in Online Resource 2 (sheet "Supplementary Table 2"). Six morphometric parameters were automatically extracted from CT-based 3D patellar reconstructions using a custom pipeline developed in 3D Slicer (version 5.10.0): patellar length, patellar width, patellar thickness, patellar volume, patellar surface area, and patellar coronal perimeter. Linear parameters are reported in millimetres (mm), volume in cubic millimetres (mm^3), and surface area in square millimetres (mm^2). All six parameters showed significant sexual dimorphism in the internal cohort (two-sided Wilcoxon rank-sum test, all $P < 0.001$).

Supplementary Table 3 Full set of radiomics features extracted from patellar regions of interest

Each row corresponds to one study participant, and each column corresponds to one radiomic feature. The full dataset is provided in Online Resource 2 (sheet "Supplementary Table 3"). Features were extracted from segmented patellar regions of interest (ROIs) on CT images using PyRadiomics (version 3.0.1). Feature classes include first-order statistics, shape features, and texture features (grey-level co-occurrence matrix, grey-level run length matrix, grey-level size zone matrix, grey-level dependence matrix, and neighbourhood grey-tone difference matrix). All feature values are presented in their original scale, covering all participants from the training, internal test and external validation sets.

Abbreviations: GLCM, grey-level co-occurrence matrix; GLRLM, grey-level run length matrix; GLSZM, grey-level size zone matrix; GLDM, grey-level dependence matrix; NGTDM, neighbourhood grey-tone difference matrix.

Supplementary Table 4 Feature importance scores for the 2D CNN and 3D CNN models

Each row corresponds to a rank-matched pair of features from the two models, sorted in descending order of importance. The full dataset is provided in Online Resource 2 (sheet “Supplementary Table 4”). Importance scores were computed via permutation importance on the final trained models; higher scores indicate greater contribution to sex classification. Feature names and corresponding importance scores are presented side by side for the 2D CNN and 3D CNN models.

Abbreviation: CNN, convolutional neural network.