



OPEN ACCESS

EDITED BY

Feng Gao,
The Sixth Affiliated Hospital of Sun Yat-sen
University, China

REVIEWED BY

Haitao Yu,
Tianjin University, China
Yi Zheng,
Jiangnan University, China

*CORRESPONDENCE

Gong Chen,
✉ 13813893873@163.com

RECEIVED 08 July 2025

ACCEPTED 13 August 2025

PUBLISHED 26 August 2025

CITATION

Wang Y, Lan T, Dou W, Chen Z, Zhang S and
Chen G (2025) Structure-guided deep
learning for back acupoint localization via
bone-measuring constraints.
Front. Physiol. 16:1662104.
doi: 10.3389/fphys.2025.1662104

COPYRIGHT

© 2025 Wang, Lan, Dou, Chen, Zhang and
Chen. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Structure-guided deep learning for back acupoint localization via bone-measuring constraints

Yulong Wang¹, Tian Lan², Wenjian Dou³, Zhi Chen⁴,
Song Zhang³ and Gong Chen^{1,3*}

¹School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing, China, ²School of Computer Science, Inner Mongolia University, Hohhot, China, ³Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, ⁴Institute of Chinese Medicine Literature, Nanjing University of Chinese Medicine, Nanjing, China

Accurate acupoint localization is crucial for the effectiveness of acupuncture and related Traditional Chinese Medicine (TCM) therapies. This study introduces a novel automated framework for recognizing back acupoints, uniquely integrating the traditional TCM bone-measuring principle with advanced deep learning for medical image analysis. The method employs an HRFormer backbone network combined with a Structure-Guided Keypoint Estimation Module (SG-KEM) and a structure-constrained loss function, ensuring anatomically consistent predictions within a standardized spatial coordinate system to improve accuracy across diverse body types. Trained and evaluated on a dataset of 430 high-resolution back images with 19 annotated acupoints, the framework achieved a normalized mean error (NME) of 0.6%, a failure rate (FR@1 cm) of 1.2%, an area under the curve (AUC) of 0.97, and a precision of 93.8%, while operating in real-time at 18 frames per second. Component analysis confirmed significant contributions: the SG-KEM module reduced the mean error by 33.3%, and the structure-constrained loss further decreased it to 0.6%. Moreover, ablation studies under challenging conditions validated the model's robustness. On the obese subset, the NME decreased from 1.5% to 0.8%, FR@1 cm dropped from 4.0% to 1.3%, and precision improved from 83.8% to 93.4%. Under illumination variation, the model achieved an NME of 0.9%, outperforming both HRFormer (1.3%) and HRFormer+SG-KEM (1.1%), with corresponding increases in AUC and precision. These findings demonstrate strong generalization across diverse clinical scenarios. Collectively, these results establish a clinically viable and computationally efficient solution for intelligent acupoint localization, supporting AI-assisted diagnosis and personalized treatment strategies within modern TCM healthcare systems.

KEYWORDS

acupoint localization, HRFormer, anatomical landmark detection, bone-measuring method, medical imaging, artificial intelligence

1 Introduction

Traditional Chinese Medicine (TCM) is a comprehensive medical system with a history spanning thousands of years and has gained widespread application worldwide through extensive clinical practice (Fung, 2009). Rooted in the theories of zang-fu organs and meridians, TCM prominently features acupuncture and massage,

which are primarily utilized for disease prevention, treatment, and alleviation of fatigue (Epstein et al., 2023). Acupuncture and massage achieve therapeutic effects and fatigue relief by stimulating specific acupoints on the human body, thereby regulating the flow of qi and blood and achieving a balance of yin and yang (Ma, 2021; Cai and Vasconcelos, 2018). Accurate acupoint localization is critical to the efficacy of acupuncture and massage therapies, as inaccuracies can directly affect treatment outcomes. Historically, acupoint identification has often relied on clinical experience, potentially leading to inconsistent treatment effects. Therefore, the development of high-precision acupoint recognition technologies holds significant promise for enhancing acupuncture accuracy, facilitating the modernization of TCM, and promoting intelligent diagnostic systems (Qi et al., 2024; Jaladat et al., 2023; Shen et al., 2024).

Furthermore, recent studies have demonstrated that accurate stimulation of specific acupoints can modulate cortical activity and brain network connectivity, revealing the neurophysiological basis of acupuncture efficacy. EEG- and fMRI-based evidence shows that acupuncture at well-localized points can regulate spectral power and functional connectivity in the brain, contributing to therapeutic effects in neurological conditions such as epilepsy and Parkinson's disease (Xue et al., 2023; Yu et al., 2018; Yu et al., 2019; Yu et al., 2024). In particular, decoding brain responses to acupuncture using EEG representation learning has laid the foundation for intelligent acupuncture-brain interfaces (Lei et al., 2023; Yu et al., 2025). These findings highlight the critical need for precise and automated acupoint localization as a prerequisite for advancing brain-targeted TCM therapies and integrative medicine systems.

In recent years, deep learning has increasingly found applications in the medical field, offering promising opportunities for the modernization of TCM (Pan et al., 2024; Zhou et al., 2024). Yang et al. proposed a deep learning-based method for back acupoint localization on weak-feature body surfaces, incorporating attention mechanisms to enhance feature representation (Yang et al., 2024). However, their method primarily relies on pixel-level intensity and lacks anatomical structure modeling, which limits its ability to adapt to individual variations in body shape or posture. Moreover, it does not introduce normalization mechanisms like bone-based coordinate systems, which are essential for physiologically consistent keypoint detection (Zhang et al., 2024). Our method explicitly addresses these challenges by incorporating a structure-guided estimation module and a TCM-inspired bone-measuring loss function, enabling better generalization and anatomical fidelity. Researchers such as Alexopoulos et al. (2023) have used deep learning for the early detection of knee osteoarthritis, while Panda et al. (2024) have applied it to lung tissue classification. Ronneberger et al. (2015) U-net model has become a staple in biomedical image segmentation, and Lee et al. (2018) have developed convolutional neural network models for the classification of dental diseases.

Several studies have showcased the potential of deep learning in acupoint recognition (Li et al., 2024). Sun et al. (2022) focused on auricular point localization by constructing a 91-keypoint dataset and applying directional normalization modules, achieving high precision in ear-based acupuncture. Wang et al. (2023) Hand acupuncture point localization method based on a dual-attention mechanism and cascade network model to localize 21 hand acupoints with excellent real-time performance. Similarly, Yuan et al. (2024) proposed the YOLOv8-ACU

framework for facial acupoint detection, incorporating lightweight ECA modules and a Slimneck structure to balance model accuracy and efficiency. While these approaches demonstrate strong performance in their respective body regions, they primarily target areas with rich local features and fixed landmarks, such as the ears, hands, and face. In contrast, the human back lacks visually salient landmarks and exhibits considerable variation in body morphology, making acupoint localization significantly more challenging (Yang et al., 2024).

While these methods offer unique advantages in their respective body regions, research focusing on back acupoints remains limited. Compared to acupoint areas like the ear or face, back acupoint recognition presents distinct challenges due to the lack of clear reference structures, a generally flat surface, and indistinct landmarks, posing significant challenges for automated localization (Kim et al., 2023; Mao et al., 2021). Nevertheless, the back hosts numerous vital back-shu points closely linked to internal organ functions, bearing irreplaceable clinical significance (Kim et al., 2023). Therefore, improving back acupoint recognition accuracy is critical for advancing intelligent diagnosis in TCM.

To overcome these limitations, we propose a novel acupoint detection method that leverages both structural and contextual knowledge. Building upon HRFormer (Yuan Y. et al., 2021), a high-resolution transformer network for dense prediction tasks, we introduce the Structure-Guided Keypoint Estimation Module (SG-KEM) to explicitly integrate osteological priors from Traditional Chinese Medicine. In addition, we design a structure-constrained loss based on bone-proportion theory (Gang et al., 2011), enabling the model to predict acupoints within a normalized anatomical coordinate system. Unlike prior methods, our approach accounts for both pixel-wise accuracy and physiological consistency, achieving high precision while maintaining robustness across individuals with varying body proportions and imaging conditions.

2 Materials and methods

2.1 Dataset

To support model training and evaluation, we utilized the publicly available DMD-BAK dataset, which addresses the lack of large-scale annotation resources for back acupoint localization. The dataset is accessible at <https://www.kaggle.com/datasets/chunzhey/dmd-bak> and contains 2,691 high-resolution JPG images of the human back. Professional Traditional Chinese Medicine (TCM) practitioners were invited to assist in both the selection of 430 representative images—based on pose diversity, image clarity, and annotation completeness—and the re-annotation of acupoint positions to ensure accuracy and clinical validity. In addition, new annotation modules were incorporated to enrich the dataset structure and facilitate subsequent model training. Each selected image includes standardized annotations for 19 back acupoints. All data remain anonymized and ethically compliant. The final subset offers a practical balance between anatomical diversity and computational feasibility, making it well suited for deep learning-based localization tasks. Table 1 lists the corresponding acupoint codes and names, and Table 2 summarizes acupoint-skeletal correlations with topologic descriptors.

TABLE 1 The correspondence between the code numbers and the acupoints.

Seq.	Type	Code	Acupoint	Index
1	DU	DU14	Dazhui	1
2	BL	BL11	Dazhu	2/3
3	BL	BL12	Fengmen	4/5
4	BL	BL13	Feishu	6/7
5	BL	BL14	Jueyinshu	8/9
6	BL	BL15	Xinshu	10/11
7	BL	BL17	Geshu	12/13
8	BL	BL18	Ganshu	14/15
9	BL	BL19	Danshu	16/17
10	BL	BL20	Pishu	18/19
11	BL	BL21	Weishu	20/21
12	BL	BL22	Sanjiaoshu	22/23
13	BL	BL23	Shenshu	24/25
14	BL	BL25	Dachangshu	26/27
15	BL	BL43	Gaohuang	28/29
16	GB	GB21	Jianjing	30/31
17	SI	SI9	Jianzhen	32/33
18	SI	SI10	Naoshu	34/35
19	SI	SI11	Tianzong	36/37

To ensure robust model evaluation, the dataset was divided into training, validation, and testing subsets at a 6:2:2 ratio. The validation set was strictly separated from the training data to prevent data leakage, allowing for accurate monitoring of generalization performance and effective hyperparameter tuning.

2.2 Data preprocessing

Given the relatively limited dataset size, multiple data augmentation techniques were applied exclusively to the training set to enhance generalization and model robustness. These included random horizontal flipping ($p = 0.5$), affine transformations (rotation within $\pm 15^\circ$, scaling between 0.9 and 1.1), and adjustments in brightness and contrast (scaling factors between 0.8 and 1.2). All augmented samples retained label consistency (Wu et al., 2022). The test set remained unaltered to ensure fair evaluation (Table 3), summarizing the augmentation strategies.

Following data cleansing and augmentation, the final dataset included 258 training images, 86 validation images, and 86 testing images, all manually verified for clarity and annotation accuracy.

2.3 Backbone network

The HRFormer architecture was adopted as the backbone of our model due to its superior performance in dense prediction tasks. HRFormer integrates the multi-resolution parallel structure of HRNet with the global modeling capability of Transformers. The network comprises four stages (Stage 1 to Stage 4), each containing multiple branches of varying resolutions. Transformer blocks within each stage operate in windowed self-attention mode, and feature maps are fused across branches to preserve both fine-grained spatial information and high-level semantic understanding.

(Table 4) outlines the structural configuration of HRFormer. Maintaining a high-resolution stream throughout the network enables the precise representation of acupoint features. Furthermore, the network's capacity to process skeletal structures (e.g., spinal curvature, scapular positions) alongside fine local details (e.g., inter-acupoint distances) enhances both accuracy and generalization. The architecture is illustrated in (Figure 1).

2.4 The SG-KEM module

To address the challenge of low visual salience in back acupoint recognition, we propose the Structure-Guided Keypoint Estimation Module (SG-KEM). This module is designed to integrate prior knowledge of human skeletal structures with context-aware features from the neighborhoods of keypoints, guiding the model to focus on anatomically meaningful regions and thereby improving the robustness and accuracy of keypoint localization. SG-KEM consists of two submodules: the Structural Prior Enhancement Module (SPEM), which models the relationship between acupoints and skeletal structures to provide structural guidance; and the Local Context Attention Module (LCAM), which enhances semantic representation in local regions through a lightweight attention mechanism. These two components work synergistically to improve the model's ability to adapt to complex backgrounds and individual variations (Figure 2).

2.4.1 Structure prior enhancement module (SPEM)

SPEM introduces anatomical constraints derived from Traditional Chinese Medicine (TCM) knowledge by leveraging a set of bone-referenced landmarks $B = \{b_k\}_{k=1}^K$, such as the seventh cervical vertebra and inferior scapular angles. These landmarks are manually annotated and exhibit stable relative positions across individuals.

We construct a set of geometrical edges Equation 1:

$$\mathcal{E} = \{(i, j) \mid b_i, b_j \in B, D_{ij}^{\text{ref}} \text{ known}\} \quad (1)$$

For each pair $(i, j) \in \mathcal{E}$, we compute a distance map (D_{ij}) and a directional map (A_{ij}) based on predicted keypoint locations (p_i, p_j) Equation 2.

$$D_{ij}(x, y) = |p_i - p_j|_2, A_{ij}(x, y) = \frac{p_i - p_j}{|p_i - p_j|_2} \quad (2)$$

These are concatenated to form a structural guidance tensor G_{prior} , which is fused with the HRFormer feature map F via a (1×1)

TABLE 2 Acupoint-skeletal correlations with topologic descriptors.

Seq.	Acupoint	Skeletal landmark	Topological relationship
1	Dazhui	C7 spinous process	Below C7 (midline)
2	Dazhu	T1 Spinous Process	Below T1, 1.5 cun lateral
3	Fengmen	T2 Spinous Process	Below T2, 1.5 cun lateral
4	Feishu	T3 Spinous Process	Below T3, 1.5 cun lateral
5	Jueyinshu	T4 Spinous Process	Below T4, 1.5 cun lateral
6	Xinshu	T5 Spinous Process	Below T5, 1.5 cun lateral
7	Geshu	T7 Spinous Process	Below T7, 1.5 cun lateral (scapula level)
8	Ganshu	T9 Spinous Process	Below T9, 1.5 cun lateral
9	Danshu	T10 Spinous Process	Below T10, 1.5 cun lateral
10	Pishu	T11 Spinous Process	Below T11, 1.5 cun lateral
11	Weishu	T12 Spinous Process	Below T12, 1.5 cun lateral
12	Sanjiaoshu	L1 Spinous Process	Below L1, 1.5 cun lateral
13	Shenshu	L2 Spinous Process	Below L2, 1.5 cun lateral (umbilicus level)
14	Dachangshu	L4 Spinous Process	Below L4, 1.5 cun lateral (iliac crest level)
15	Gaohuang	T4 Spinous Process	Below T4, 3 cun lateral (medial scapula)
16	Jianjing	C7 and Acromion	Midpoint of C7-Acromion line
17	Jianzhen	Posterior Axillary Fold	1 cun above axillary fold
18	Naoshu	Scapular Spine	Inferior to scapular spine
19	Tianzong	Scapular Spine	Center of infraspinous fossa

TABLE 3 Data augmentation strategies.

Augmentation type	Description	Parameter range/Probability
Random Horizontal Flip	Symmetric left-right flip of the image	Probability = 0.5
Random Rotation	Rotation operation in affine transformation	Angle ∈ [−15°, +15°]
Random Scaling	Affine scaling operation with the center unchanged	Scaling ratio ∈ [0.9, 1.1]
Random Brightness Perturbation	Adjust the overall brightness of the image	Adjustment factor ∈ [0.8, 1.2]
Random Contrast Perturbation	Change the contrast of the image	Adjustment factor ∈ [0.8, 1.2]

convolution **Equation 3**:

$$F_{SPEM} = F + Conv_{1 \times 1} \left(Concat_{(i,j) \in \mathcal{E}} \left[D_{ij}, A_{ij}\right]\right)$$

(3)

This fusion enables the network to be aware of physiologically plausible acupoint arrangements.

2.4.2 Local contextual attention module (LCAM)

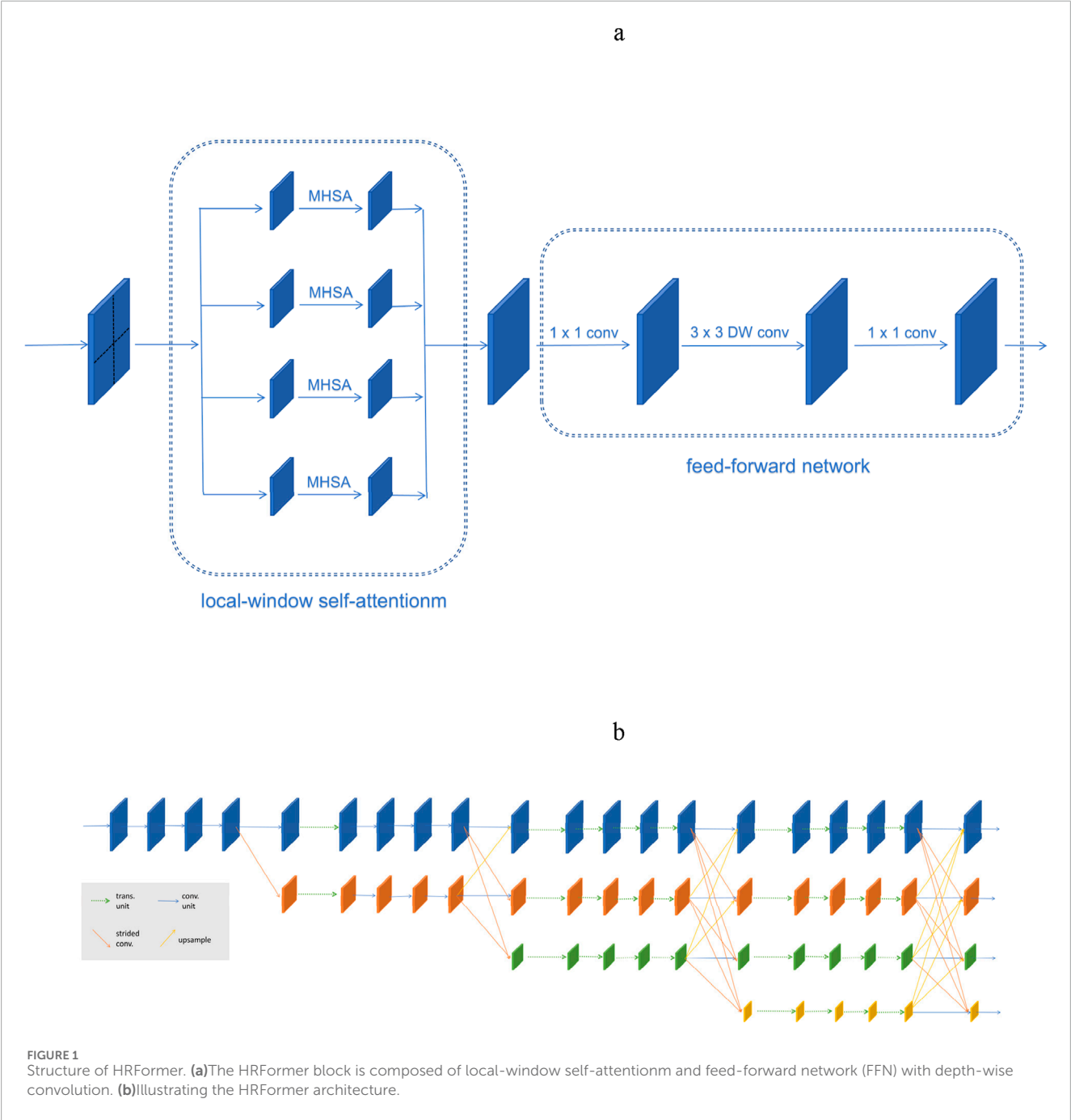
To model local dependencies and eliminate background noise, LCAM applies directional convolution and spatial attention. The SPEM-enhanced feature F_{SPEM} is first processed with decomposed directional convolutions **Equation 4**:

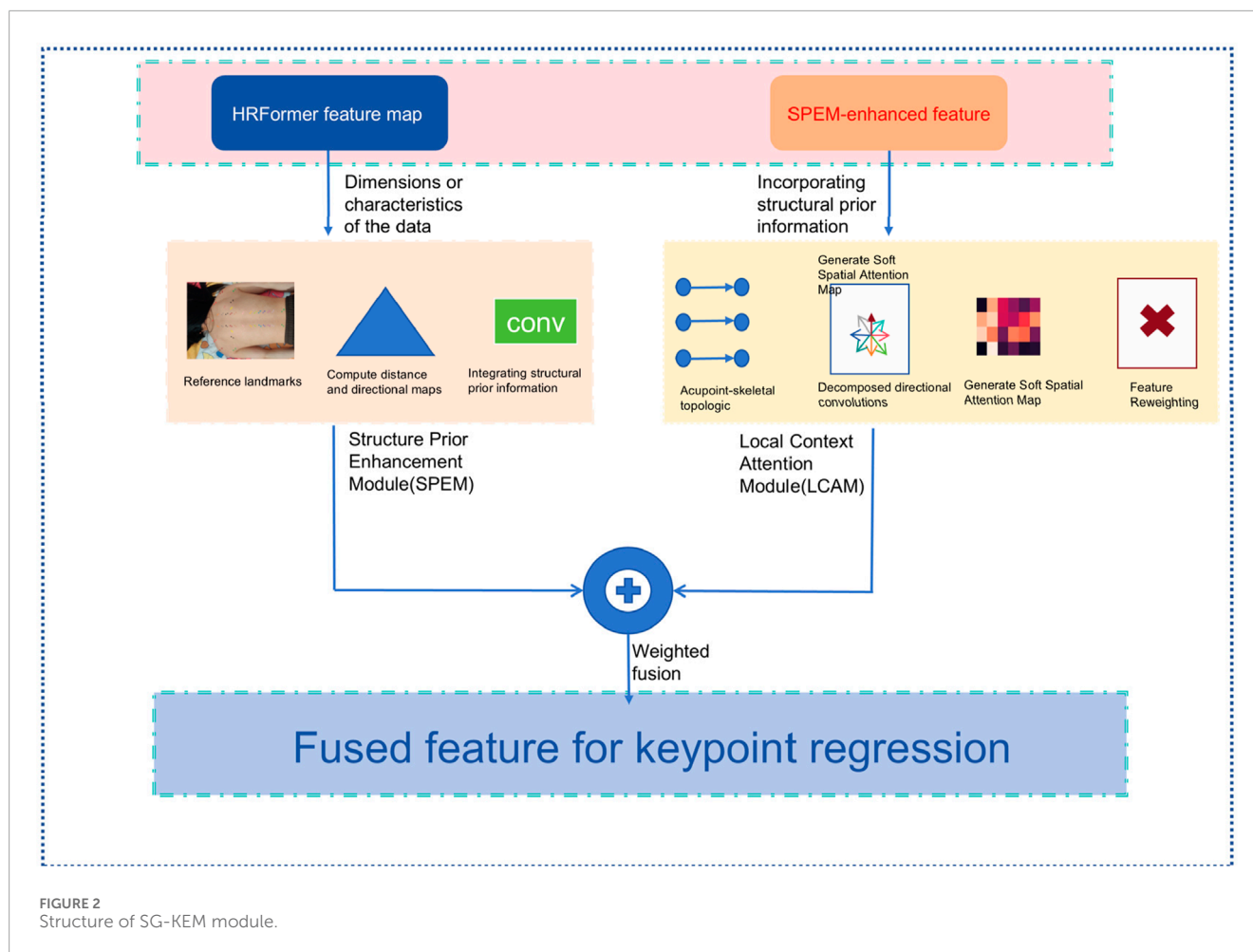
$$F_{dir} = ReLU(Conv_{3 \times 1}(F_{LCAM}) + Conv_{1 \times 3}(F_{SPEM}))$$

(4)

TABLE 4 Parameters of four stages of the HRFormer.

Stage	Number of branches	Resolution ratio	Number of transformer blocks per branch	Embedding dimension per layer(C)
1	1	1/4	2	64
2	2	1/4,1/8	2,2	64,128
3	3	1/4,1/8,1/16	2,2,2	64,128,256
4	4	1/4,1/8,1/16,1/32	2,2,2,2	64,128,256,512





Next, we generate a soft spatial attention map Equation 5:

$$A_{(x,y)} = \frac{\exp(Q_{(x,y)})}{\sum_{(u,v)} \exp(Q_{(u,v)})}, Q = \text{Conv}_{1 \times 1} F_{dir} \quad (5)$$

The final output is a reweighted feature map Equation 6:

$$F_{LCAM} = F_{dir} \odot A \quad (6)$$

2.4.3 Output fusion

We fuse the outputs of SPEM and LCAM to obtain the final structurally enhanced feature representation for keypoint regression Equation 7:

$$F_{SG-KEM} = (\text{Conv}_{3 \times 3} (F_{SPEM} + F_{LCAM})) \quad (7)$$

This fused representation is passed to the keypoint regression head for heatmap generation.

2.5 Structure-constrained loss function

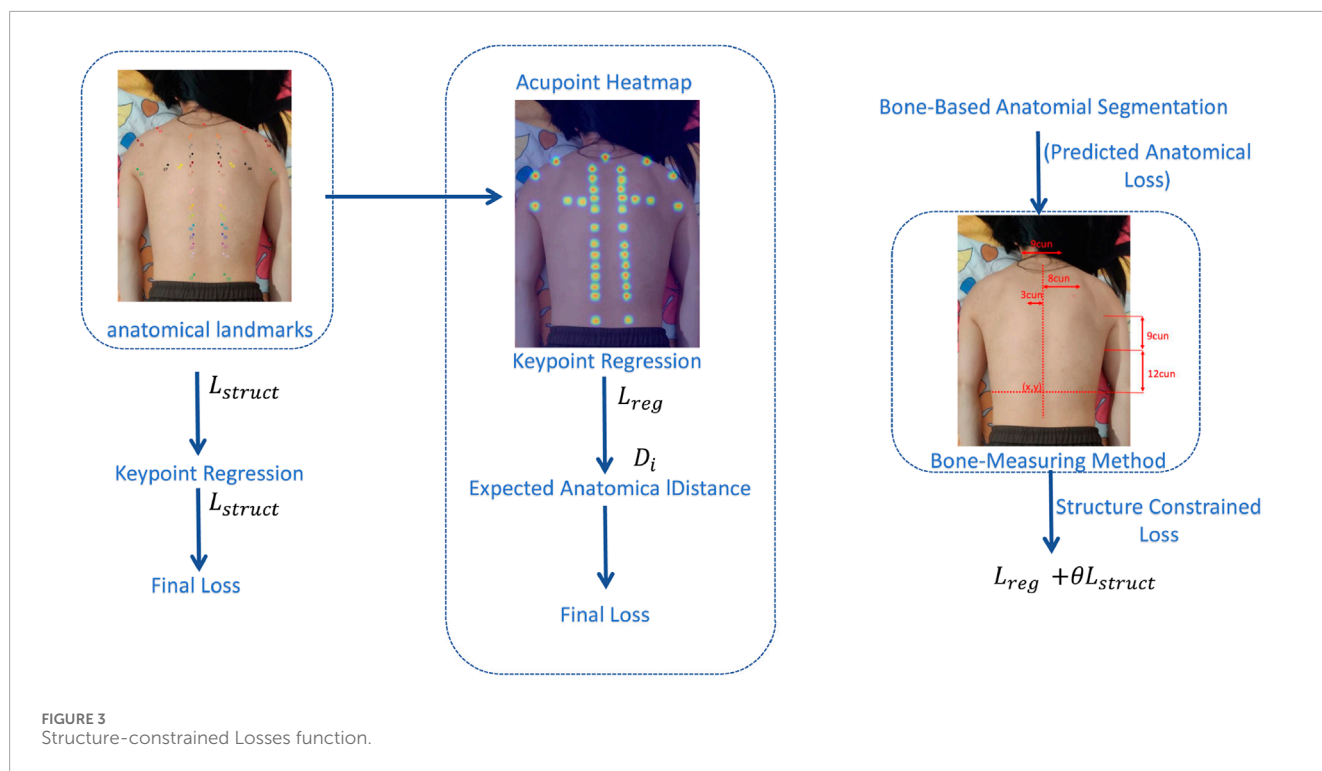
To fully leverage the structure-guided features extracted by the SG-KEM module, the resulting feature map is used as the

output of the keypoint heatmap branch to regress the spatial coordinates of acupoints. To enhance both localization accuracy and anatomical plausibility, we design a structure-constrained loss function based on the traditional Chinese medicine (TCM) bone-measuring method.

Unlike general pose estimation tasks involving full-body joint detection, our method only focuses on stable anatomical landmarks in the back region relevant to acupoint localization. These landmarks obtained through a lightweight anatomical landmark detection module. Based on these reference points, we divide the trunk region into standardized proportional units known as “cuns” using the TCM bone-measuring method. This process constructs a subject-specific proportional reference space, onto which all acupoint annotations are projected based on proportional units (cun).

The normalized coordinate space eliminates differences in body proportion and posture, allowing the model to learn acupoint localization in a structurally consistent and interpretable manner. The architectural framework of this method is shown in (Figure 3).

To incorporate anatomical structure into the training process, we introduce a sample-specific normalization mechanism. Specifically, we estimate a personalized scale factor r , representing the number of pixels per cun for each sample. This factor is computed using the Euclidean distance d between two stable anatomical landmarks—typically the medial borders of the left



and right scapulae—defined to span 24 cuns according to TCM standards [Equation 8](#):

$$r = \frac{d}{24} \quad (8)$$

This normalization enables the model to apply structural constraints in a physiologically consistent coordinate space, allowing adaptation to individuals with varying body proportions while preserving standard acupoint relationships.

To encode structural knowledge, we define a set of acupoint pairs \mathcal{P} with known anatomical distances derived from TCM literature. These include bilateral Dachangshu (BL25), Xinshu (BL15), and spinal-axis acupoints such as Dazhui (DU14). For each acupoint pair, the expected distance in cuns d_{cun} is translated into pixel space using the personalized scale α .

According to TCM standards, the distance between specific acupoint pairs (e.g., bilateral BL25) is defined as a fixed number of cuns (e.g., six cuns), which is then converted to pixel distance by multiplying with the scale factor α [Equation 9](#):

$$d_{pix}^{expected} = \alpha \cdot d_{cun} \quad (9)$$

This step ensures anatomical distances are expressed in the same coordinate system as the predicted keypoint locations, enabling consistent comparison during optimization.

The structure-constrained loss is formulated as a mean squared error between the predicted distance and the expected anatomical distance [Equation 10](#):

$$\mathcal{L}_{struct} = \frac{1}{M} \sum_{(i,j) \in \mathcal{P}} \left(\|\hat{p}_i - \hat{p}_j\|_2 - r \cdot l_{ij} \right)^2 \quad (10)$$

Here, \mathcal{P} denotes the set of acupoint pairs selected for structural constraint, and M is the total number of constrained pairs.

To jointly optimize both pixel-level accuracy and anatomical consistency, we define the final loss function as a weighted sum of keypoint regression loss \mathcal{L}_{reg} and the structural loss \mathcal{L}_{struct} [Equation 11](#):

$$\mathcal{L}_{total} = \mathcal{L}_{reg} + \theta \cdot \mathcal{L}_{struct} \quad (11)$$

Where θ is a hyperparameter balancing precision and structural compliance.

2.6 Evaluation metrics

To evaluate the performance of the proposed model, we adopted five standard metrics: Normalized Mean Error (NME) ([Lai et al., 2019](#)), Failure Rate (FR) ([Finkelstein, 2008](#)), Area Under the Curve (AUC) ([Myerson et al., 2001](#)), Precision ([Streiner and Norman, 2006](#)), and Frames Per Second (Image/s) ([Koslowsky et al., 2006](#)). These metrics jointly assess the model's accuracy, robustness, and efficiency.

Normalized Mean Error (NME): The average Euclidean distance between predicted and ground-truth keypoints, normalized by inter-Xinshu distance [Equation 12](#):

$$NME = \frac{1}{K \cdot N} \sum_{k=1}^K \sum_{n=1}^N \frac{\|Y_k^n - \hat{Y}_k^n\|}{d} \times 100\% \quad (12)$$

Failure Rate (FR): The proportion of test samples with NME exceeding a fixed threshold δ [Equation 13](#):

$$FR = \frac{1}{K} \sum_{k=1}^K \mathbb{I} \left(\frac{1}{N} \sum_{n=1}^N \frac{\|Y_k^n - \hat{Y}_k^n\|}{d} > \delta \right) \times 100\% \quad (13)$$

Area Under the Curve (AUC): The integral of the cumulative error distribution curve from 0 to δ Equation 14:

$$AUC = \int_0^{10\%} CED(\delta) d\delta \quad (14)$$

Precision: The proportion of true positive predictions among all positive predictions Equation 15:

$$Precision = \frac{TP}{TP + FP} \quad (15)$$

Images Per Second (IPS): The number of Images per second Equation 16:

$$IPS = \frac{Images}{s} \quad (16)$$

3 Result

The proposed model was trained and evaluated in a high-performance computing environment equipped with an NVIDIA GeForce RTX 3080 Ti GPU, running Ubuntu 18.04. The implementation was based on Python 3.8 and PyTorch 1.10. The Adam optimizer was used, with a batch size of 32 and a dropout rate of 0.5 to mitigate overfitting. Training was conducted over 100 epochs with early stopping based on validation performance. These configurations ensured training efficiency and model generalization.

3.1 Performance comparison with baseline models

To evaluate the effectiveness of the proposed method, we conducted comparative experiments against several mainstream models, including Vision Transformer (ViT) (Yuan L. et al., 2021), ACFormer (Zong et al., 2023), RTMpose (Jiang et al., 2023; He et al., 2024), Faster R-CNN (Bharati and Pramanik, 2019), YOLOv8 (Sohan et al., 2024), HRFormer (Yuan Y. et al., 2021) and Uniformer (Li et al., 2023). The comparison focused on five key metrics: Normalized Mean Error (NME), Failure Rate within 1 cm (FR@1 cm), Area Under the Curve (AUC), Precision, and Images Per Second (IPS). The results are presented in (Figure 4).

The Normalized Mean Error (NME) dropped significantly to 0.6%, representing a 57.1% reduction compared to ViT (2.8%) and a 50% reduction compared to HRFormer (1.2%). Similarly, the FR@1 cm metric improved markedly, reaching only 1.2%, versus 12.5% for ViT and 3.8% for HRFormer, indicating high localization precision at the anatomical level.

In terms of AUC, the proposed method achieved 0.97, outperforming all others, including Uniformer (0.94) and HRFormer (0.92). Precision reached 93.8%, compared to 83.6% for YOLOv8 and 84.1% for HRFormer. Despite being slightly slower than YOLOv8 (18 FPS vs 45 FPS), the proposed model still meets real-time requirements and delivers significantly higher accuracy. These results suggest a well-balanced trade-off between speed and accuracy, with clear superiority in anatomical alignment and model generalization.

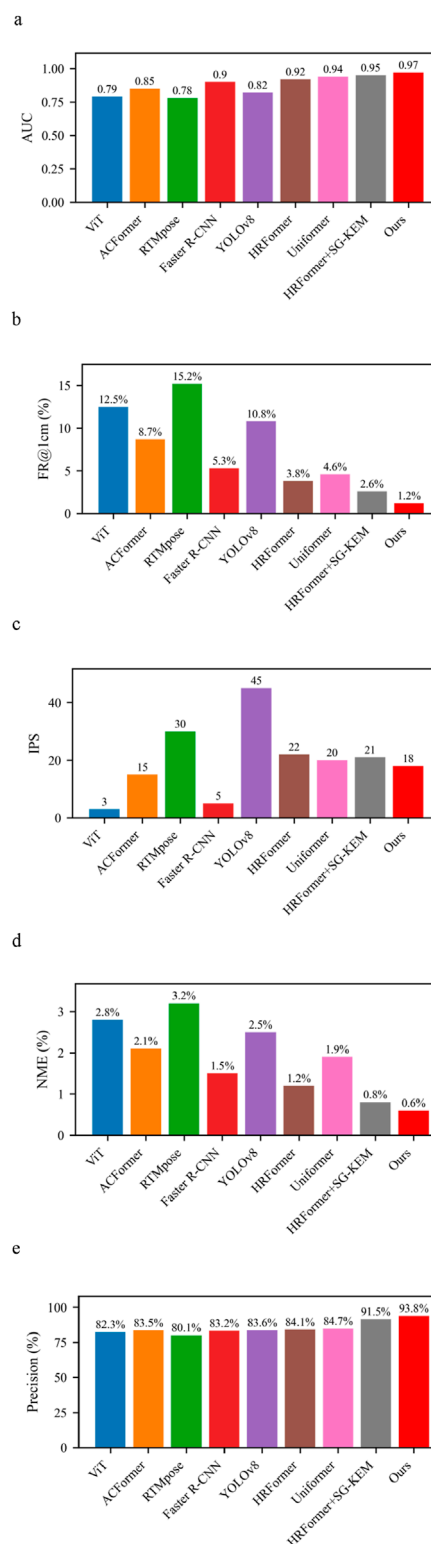


FIGURE 4
Performance comparison of different models. (a) Comparison of different models in AUC. (b) Comparison of different models in FR@1 cm (%). (c) Comparison of different models in IPS. (d) Comparison of different models in NME (%). (e) Comparison of different models in Precision (%).

These results underscore that while high-speed detectors like YOLOv8 may be suitable for general object detection, their precision is insufficient for delicate clinical tasks such as acupoint localization. Our model's structure-guided design ensures not only numerical superiority but also anatomical plausibility, essential for safe acupuncture guidance or medical navigation.

3.2 Robustness under obese body morphologies

To assess the robustness of our model in the presence of anatomical variations, we evaluated performance on a subset of obese subjects, whose back contours exhibit significant curvature, skin folds, and less-defined anatomical landmarks (Figure 5).reports the results of an ablation study under these challenging conditions.

We compared three variants:

HRFormer baseline, HRFormer + SG-KEM module and Full model (HRFormer + SG-KEM + structure-constrained loss).

NME decreased from 1.5% to 0.8%, representing a 46.7% improvement, while FR@1 cm dropped from 4.0% to 1.3%, confirming improved localization accuracy in anatomically complex regions. The addition of SG-KEM alone already improved NME to 1.3%, indicating that multi-scale structural priors provide benefit even without the loss constraint.

AUC improved from 0.91 to 0.95, and Precision increased from 83.8% to 93.4%, supporting the effectiveness of structure-constrained learning in handling size-induced anatomical distortion. FPS declined modestly from 22 to 18, but the real-time capability remained acceptable. These results indicate that our structure-guided model is well-adapted to variations in body morphology, enhancing its clinical applicability for diverse patient groups.

The improvements in obese individuals are particularly significant from a clinical perspective. In practice, acupoint palpation in overweight patients is more difficult due to tissue coverage and ambiguous bone landmarks. By leveraging a structure-constrained spatial normalization (TCM bone-measuring system), our model achieves robust predictions across body types. This is critical for real-world deployment in diverse populations, such as in hospitals or mobile healthcare units.

3.3 Generalization under illumination variation

Lighting conditions significantly impact image-based recognition systems, particularly in clinical settings where consistent lighting is hard to maintain. To evaluate generalization under such visual disturbances, we applied the same three model variants (HRFormer, HRFormer+SG-KEM, full model) to a test set modified with varying brightness and contrast levels. Results are summarized in (Figure 6).

The NME reduced from 1.3% in HRFormer to 0.9% in the full model, a 30.8% improvement, while the FR@1 cm dropped from 4.3% to 1.5%. These findings show the proposed model's resilience in retaining spatial accuracy under degraded visual conditions.

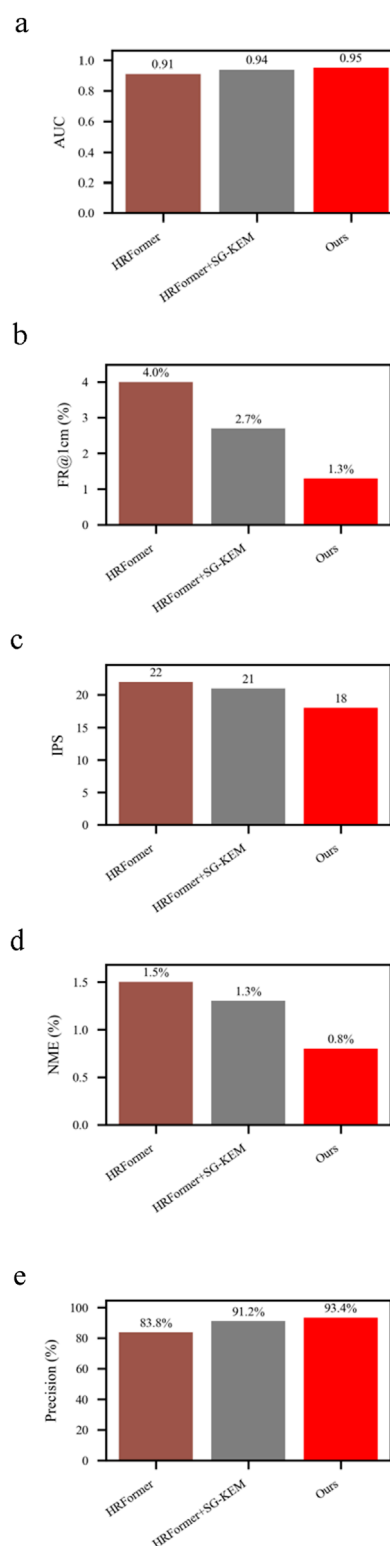


FIGURE 5
Results of ablation study under the obese subset. (a) Comparison of different models in AUC. (b) Comparison of different models in FR@1 cm (%). (c) Comparison of different models in IPS. (d) Comparison of different models in NME (%). (e) Comparison of different models in Precision (%).

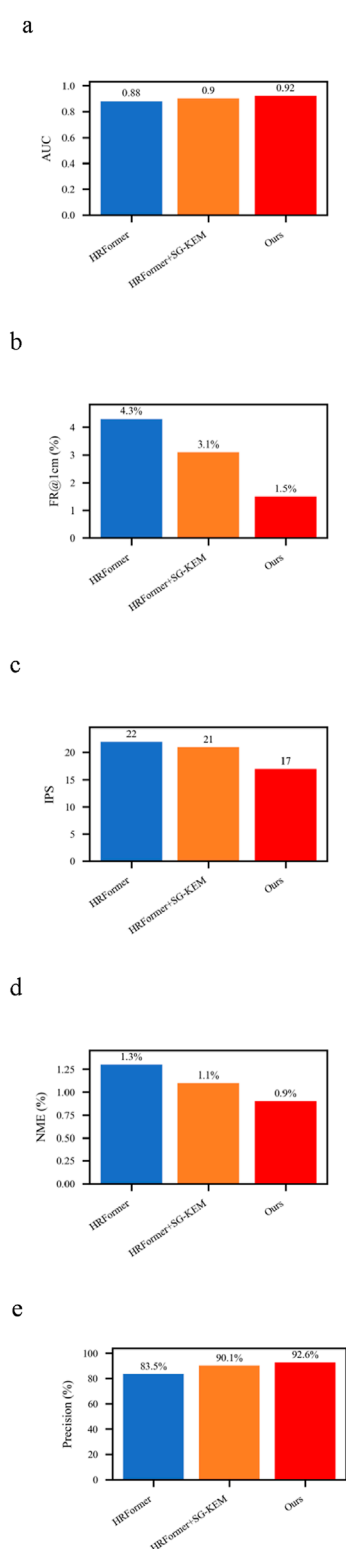


FIGURE 6
Results of ablation study under illumination variation. (a) Comparison of different models in AUC. (b) Comparison of different models in FR@1 cm (%). (c) Comparison of different models in IPS. (d) Comparison of different models in NME (%). (e) Comparison of different models in Precision (%).

AUC increased from 0.88 to 0.92, and Precision improved from 83.5% to 92.6%, confirming that both SG-KEM and structure-aware supervision enhance the model's lighting invariance. Although FPS slightly decreased from 22 to 17, the model maintained real-time performance. These outcomes demonstrate that anatomical constraints improve illumination robustness, making the model suitable for dynamic clinical environments where lighting conditions may be inconsistent.

These results highlight that illumination invariance is not solely a function of network depth or capacity, but greatly benefits from domain-specific anatomical priors. The model's integration of bone-referenced geometry enables it to rely less on pixel intensity and more on structural layout, a desirable trait in noisy or uncontrolled environments. This is particularly relevant for real-world TCM diagnosis settings using mobile devices or home-care robotics.

3.4 Summarize

Across all test conditions—standard, obese, and illumination-perturbed—the proposed model consistently demonstrates superior accuracy, robustness, and clinical relevance. The improvements observed in (Figures 4–6) stem from carefully designed modules that bridge anatomical priors with deep feature extraction. This structure-aware strategy enables robust acupoint localization suitable for complex and variable clinical environments.

4 Discussion

This study proposes a structure-aware acupoint localization framework that effectively integrates Traditional Chinese Medicine (TCM) principles with advanced deep learning techniques (Gang et al., 2011). By incorporating a high-resolution transformer backbone (HRFormer), a Structure-Guided Keypoint Estimation Module (SG-KEM), and a structure-constrained loss based on the bone-measuring method, the model achieves accurate and anatomically coherent localization of back acupoints.

The integration of SG-KEM significantly enhances spatial feature representation by guiding the model to focus on physiologically meaningful regions. This module leverages skeletal priors—such as scapular and spinal landmarks—that remain relatively stable across individuals, enabling the model to localize acupoints accurately even in anatomically ambiguous or low-contrast regions (Riegler et al., 2015). Furthermore, the structure-constrained loss enforces consistency in relative acupoint spacing based on TCM-defined proportions, enhancing physiological plausibility and improving generalization across different body types.

Experimental results confirm the effectiveness of the proposed approach. On a dataset comprising 430 back images with 19 annotated acupoints, the model achieved a normalized mean error (NME) of 0.6%, a failure rate (FR@1 cm) of 1.2%, and

an AUC of 0.97. These outcomes reflect substantial gains in both spatial precision and anatomical consistency compared to multiple baselines, including HRFormer (Yuan Y. et al., 2021) alone, ViT (Yuan L. et al., 2021), and YOLOv8 (Sohan et al., 2024). Importantly, the model maintained real-time inference capability (18 IPS), meeting the demands of clinical deployment scenarios such as intelligent diagnosis, robot-assisted acupuncture, and posture-adaptive therapy systems (Cheng et al., 2024).

Ablation studies demonstrate the individual contributions of each component. The SG-KEM module reduced NME by 33.3%, while the addition of the structure-constrained loss further improved anatomical compliance and lowered NME to 0.6%. The benefits of this architecture extended to challenging clinical conditions. Under the obese subset (Mao et al., 2021), the proposed method achieved a 46.7% relative reduction in NME compared to HRFormer (from 1.5% to 0.8%), while improving precision from 83.8% to 93.4%. Similarly, under illumination variation, the model's NME decreased from 1.3% to 0.9%, and precision increased from 83.5% to 92.6%, outperforming both the backbone and intermediate variants. These results substantiate the model's robustness to patient morphology, lighting inconsistencies, and anatomical complexity, underscoring its viability for real-world clinical applications where such variability is common. This modular design supports flexible adaptation to various clinical applications and suggests potential for extension to other TCM body regions. These results substantiate the model's robustness to patient morphology, lighting inconsistencies, and anatomical complexity, underscoring its viability for real-world clinical applications where such variability is common. In particular, the framework achieved consistent accuracy in images of obese individuals, where back curvature and surface texture differ significantly from standard anatomical presentations. These findings highlight the model's resilience to intra-population variability and its suitability for broader clinical use.

Moreover, the proposed framework exemplifies how TCM domain knowledge can be systematically encoded into modern AI systems to enhance interpretability and trustworthiness. By embedding fixed anatomical constraints within both the network structure and training objective, the model yields acupoint predictions that are not only numerically accurate but also aligned with clinical and diagnostic expectations. This structure-aware paradigm offers a valuable reference for future development of AI systems that bridge traditional medical expertise with data-driven methodologies.

These findings not only confirm the technical robustness of the model but also support its alignment with classical meridian theories in Traditional Chinese Medicine (Yang et al., 2011). The accurate mapping of acupoints in proportionally normalized anatomical space resonates with the TCM concept of "bone-based measurement", bridging empirical knowledge with data-driven precision. This provides a valuable foundation for modernizing diagnostic protocols and ensuring consistency in acupuncture-based interventions across practitioners and institutions.

While the current framework demonstrates robust performance on a curated dataset of back images, its application to dynamic scenarios—such as real-time tracking during respiration or movement—remains to be explored. Additionally, future work

may incorporate multimodal sensing, including depth, thermal, or surface EMG signals, to further enhance performance under occlusion, poor lighting, or patient movement. Deployment optimization through model compression techniques may also broaden accessibility to portable and embedded hardware platforms (Pan et al., 2024; Dantas et al., 2024).

In summary, this study presents a clinically viable, structurally grounded, and computationally efficient solution for back acupoint localization. It underscores the potential of combining TCM anatomical principles with high-resolution deep learning to advance intelligent diagnosis and personalized treatment in integrative medicine.

5 Conclusion

In this study, we developed a novel acupoint localization framework that effectively integrates Traditional Chinese Medicine (TCM) anatomical knowledge with high-resolution deep learning techniques. By embedding the bone-measuring method into both the feature extraction and optimization stages, the proposed model achieves accurate, anatomically consistent localization of back acupoints—a task traditionally hindered by the back's flat morphology and sparse visual landmarks.

Our approach combines the HRFormer backbone with a Structure-Guided Keypoint Estimation Module (SG-KEM) and a structure-constrained loss function rooted in TCM's proportional anatomy. This design enables the model to capture spatially meaningful features and maintain physiologically plausible acupoint arrangements across individuals with varying body types. Experimental results demonstrate excellent localization performance (NME: 0.6%, FR@1 cm: 1.2%, AUC: 0.97), robust generalization under challenging conditions such as obesity and illumination variation, and real-time inference capability (18 IPS), confirming the model's potential for clinical deployment. Importantly, ablation studies further revealed the model's strong generalization under challenging conditions. Under the obese subset, the framework reduced NME from 1.5% to 0.8% and improved precision from 83.8% to 93.4%. In illumination variation scenarios, it achieved a 30.8% relative NME reduction (from 1.3% to 0.9%) and maintained high precision at 92.6%. These results confirm the model's robustness across diverse body shapes and imaging environments, which are commonly encountered in real-world clinical settings.

Beyond technical contributions, this work exemplifies a promising direction in AI-assisted integrative medicine: translating traditional anatomical systems into machine-understandable priors. The framework offers a scalable and interpretable foundation for intelligent acupuncture navigation, standardized treatment planning, and TCM digitization. In the broader context of TCM, our method lays the groundwork for standardizing acupoint localization in clinical acupuncture, facilitating the integration of AI into meridian-based therapies. Furthermore, by quantifying traditionally qualitative anatomical concepts, it contributes to the digital transformation of acupuncture education, robot-assisted therapy, and international standard formulation. Future research will focus on expanding the dataset to include dynamic scenarios and diverse populations, incorporating multimodal imaging, and

optimizing the model for deployment on portable diagnostic or therapeutic devices.

Data availability statement

The datasets [DMD-BAK] for this study can be found in the [<https://www.kaggle.com/datasets/chunzheye/dmd-bak>].

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

YW: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Visualization, Writing – original draft, Writing – review and editing. TL: Project administration, Software, Supervision, Writing – original draft, Writing – review and editing. WD: Software, Visualization, Writing – original draft, Writing – review and editing. ZC: Formal Analysis, Validation, Writing – original draft. SZ: Data curation, Writing – original draft. GC: Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was

funded by 2024 Jiangsu Province Research and Practice Innovation Program Project, grant number KYCX24_2159 and the 2023 Nanjing Major Science and Technology Project (Life and Health), grant number 202305036.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2025.1662104/full#supplementary-material>

References

- Alexopoulos, A., Hirvasniemi, J., Klein, S., Donkervoort, C., Oei, E. H. G., and Tümer, N. (2023). Early detection of knee osteoarthritis using deep learning on knee magnetic resonance images. *arXiv Prepr. arXiv:2209.01192* 3, 100112. doi:10.1016/j.ostima.2023.100112
- Bharati, P., and Pramanik, A. (2019). "Deep learning techniques—R-CNN to mask R-CNN: a survey," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 6154–6162. doi:10.1109/CVPR.2018.00644
- Cai, Z., and Vasconcelos, N. (2018). "Cascade R-CNN: delving into high quality object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 6154–6162. doi:10.1109/CVPR.2018.00644
- Cheng, X., Shen, Z., and Zhang, Y. (2024). Bioinspired 3D flexible devices and functional systems. *Natl. Sci. Rev.* 11 (3), nwad314. doi:10.1093/nsr/nwad314
- Dantas, P. V., da Silva Jr, W. S., Cordeiro, L. C., and Carvalho, C. B. (2024). A comprehensive review of model compression techniques in machine learning. *Appl. Intell.* 54 (22), 11804–11844. doi:10.1007/s10489-024-05747-w
- Epstein, A. S., Zhou, J., Lee, J. Y., Carino, G., Chang, B. H., Prigerson, H. G., et al. (2023). Acupuncture vs massage for pain in patients living with advanced cancer: the IMPACT randomized clinical trial. *JAMA Netw. Open* 6 (11), e2342482. doi:10.1001/jamanetworkopen.2023.42482
- Finkelstein, M. (2008). *Failure rate modelling for reliability and risk*. Berlin: Springer. doi:10.1007/978-1-84800-986-8_5
- Fung, P. C. W. (2009). Probing the mystery of Chinese medicine Meridian channels with special emphasis on the connective tissue interstitial fluid system, mechanotransduction, cells diastasis and mast cell degranulation. *Chin. Med.* 4 (1), 10. doi:10.1186/1749-8546-4-10
- Gang, W., Yang, L., Wu, X., Zi, M., Wang, Y., and Yang, J. (2011). Study on the sequence in formulating standards for acupuncture and moxibustion. *J. Traditional Chin. Med.* 31 (2), 136–140. doi:10.1016/S0254-6272(11)60028-9
- He, S., Wei, M., Meng, D., Lv, Z., Guo, H., Yang, G., et al. (2024). Adversarially trained RTMPose: a high-performance, non-contact method for detecting genu valgum in adolescents. *Comput. Biol. Med.* 183, 109214. doi:10.1016/j.compbiomed.2024.109214
- Jaladat, A. M., Rezaeizadeh, A. R., Ansari, M. R., Mokhtari, H., Moadab, M. H., Fereidooni, H., et al. (2023). Similarities and differences between kiai in Persian medicine and moxibustion in Chinese medicine. *J. Integr. Med.* 21 (4), 354–360. doi:10.1016/j.joim.2023.05.002

- Jiang, T., Lu, P., Zhang, L., Ma, N., Han, R., Lyu, C., et al. (2023). RTMpose: Real-time multi-person pose estimation based on MMPose. *arXiv Prepr. arXiv:2303.07399*. doi:10.48550/arXiv.2303.07399
- Kim, G., Kim, D., Moon, H., Yoon, D.-E., Lee, S., Ko, S.-J., et al. (2023). Acupuncture and acupoints for low back pain: systematic review and meta-analysis. *Am. J. Chin. Med.* 51 (2), 223–247. doi:10.1142/S0192415X23500131
- Koslowsky, B., Jacob, H., Eliakim, R., and Adler, S. N. (2006). PillCam ESO in esophageal studies: improved diagnostic yield of 14 frames per second (fps) compared with 4 fps. *Endoscopy* 38 (1), 27–30. doi:10.1055/s-2005-921034
- Lai, S., Chai, Z., Li, S., Meng, H., Yang, M., and Wei, X. (2019). “Enhanced normalized mean error loss for robust facial landmark detection,” in *Proceedings of the british machine vision conference (BMVC)*. doi:10.5244/C.33.87
- Lee, J.-H., Kim, D.-H., Jeong, S.-N., and Choi, S.-H. (2018). Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.* 77, 106–111. doi:10.1016/j.jdent.2018.07.015
- Lei, S., Fan, J., Liu, X., Xu, X., Zhang, J., Zhou, Z., et al. (2023). Qualitative and quantitative meta-analysis of acupuncture effects on the motor function of Parkinson's disease patients. *Front. Neurosci.* 17, 1125626. doi:10.3389/fnins.2023.1125626
- Li, K., Wang, Y., Zhang, J., Peng, G., Song, G., Liu, Y., et al. (2023). UniFormer: unifying convolution and self-attention for visual recognition. *IEEE Trans. Pattern Analysis Mach. Intell.* 45 (10), 12581–12600. doi:10.1109/TPAMI.2023.3282631
- Li, Y., Teng, Y., Huang, Y., Huang, L., Yang, S., Liu, J., et al. (2024). AIR-Net: acupoint image registration network for automatic acupoint recognition and localization. *Displays* 83, 102743. doi:10.1016/j.displa.2024.102743
- Ma, H. (2021). “Automatic positioning system of medical service robot based on binocular vision,” in *Proceedings of the international symposium on robotics and intelligent manufacturing technology (ISRIMT)* (IEEE), 52–55. doi:10.1109/ISRIMT53730.2021.9597049
- Mao, J., Zhu, B., and Wang, H. (2021). Acupuncture-based treatments on cervical spondylosis of vertebral artery type. *TMR Non-Drug Ther.* 4 (1), 6–12. doi:10.53388/tmrnd20210303027
- Myerson, J., Green, L., and Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *J. Exp. Analysis Behav.* 76 (2), 235–243. doi:10.1901/jeab.2001.76-235
- Pan, D., Guo, Y., Fan, Y., and Wan, H. (2024). Development and application of traditional Chinese medicine using AI machine learning and deep learning strategies. *Am. J. Chin. Med.* 52 (3), 605–623. doi:10.1142/S0192415X24500265
- Panda, S. K., Ramesh, J. V. N., Ghosh, H., Rahat, I. S., Sobur, A., Bijoy, M. H., et al. (2024). Deep learning in medical imaging: a case study on lung tissue classification. *EAI Endorsed Trans. Pervasive Health Technol.*, 10. doi:10.4108/eetpht.10.5549
- Qi, W., He, B., Gu, Q., Li, Y., and Liang, F. (2024). Scientific exploration and hypotheses concerning the meridian system in traditional Chinese medicine. *Acupunct. Herb. Med.* 4 (3), 283–289. doi:10.1097/HM9.0000000000000128
- Riegler, G., Urschler, M., Rüther, M., Bischof, H., and Stern, D. (2015). “Anatomical landmark detection in medical applications driven by synthetic data,” in *Proceedings of the IEEE international conference on computer vision workshops (ICCVW)* (Santiago, Chile), 12–16. doi:10.1109/ICCVW.2015.21
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: convolutional networks for biomedical image segmentation. *Med. Image Comput. Computer-Assisted Intervention (MICCAI)* 9351, 234–241. doi:10.1007/978-3-319-24574-4_28
- Shen, Y., Wang, S., Shen, Y., Xing, H., Gong, L., and Hu, J. (2024). KOA massage robot: a study on the reduction of TCM manipulation based on PSO-BP algorithm. *IEEE Access* 12, 149367–149380. doi:10.1109/ACCESS.2024.3471889
- Sohan, M., Ram, T. S., and Reddy, C. V. R. (2024). “A review on YOLOv8 and its advancements,” in *International conference on data intelligence and cognitive informatics* (Springer), 44–52. doi:10.1007/978-981-99-7962-2_39
- Streiner, D., and Norman, G. R. (2006). Precision and accuracy: two terms that are neither. *J. Clin. Epidemiol.* 59 (4), 327–330. doi:10.1016/j.jclinepi.2005.09.005
- Sun, X., Dong, J., Li, Q., Lu, D., and Yuan, Z. (2022). Deep learning-based auricular point localization for auriculotherapy. *IEEE Access* 10, 112898–112908. doi:10.1109/ACCESS.2022.3215138
- Wang, H., Liu, L., Wang, Y., and Du, S. (2023). Hand acupuncture point localization method based on a dual-attention mechanism and cascade network model. *Biomed. Opt. Express* 14, 5965–5978. doi:10.1364/BOE.501663
- Wu, X., Huang, L., and Zhao, J. (2022). Interpretation of China national standard nomenclature and location of meridian points (GB/T 12346-2021). *Chin. Acupunct. Moxibustion* 42 (5), 579–582. doi:10.13703/j.0255-2930.20220117-k0001
- Xue, H., Zeng, L., He, H., Xu, D., and Ren, K. (2023). Effectiveness of acupuncture as auxiliary combined with Western medicine for epilepsy: a systematic review and meta-analysis. *Front. Neurosci.* 17, 1203231. doi:10.3389/fnins.2023.1203231
- Yang, E. S., Li, P. W., Niluis, B., and Li, G. (2011). Ancient Chinese medicine and mechanistic evidence of acupuncture physiology. *Pflügers Arch. - Eur. J. Physiol.* 462 (5), 645–653. doi:10.1007/s00424-011-1017-3
- Yang, S., Li, Y., Zou, H., Huang, L., Liu, J., Teng, Y., et al. (2024). Exploring an innovative deep learning solution for acupuncture point localization on the weak feature body surface of the human back. *IEEE J. Biomed. Health Inf.* 29, 4599–4611. doi:10.1109/JBHI.2024.3511128
- Yu, H., Wu, X., Cai, L., Deng, B., and Wang, J. (2018). Modulation of spectral power and functional connectivity in human brain by acupuncture stimulation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26 (5), 977–986. doi:10.1109/TNSRE.2018.2828143
- Yu, H., Wang, L., Lei, X., and Wang, J. (2019). Modulation effect of acupuncture on functional brain networks and classification of its manipulation with EEG signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27 (10), 1973–1984. doi:10.1109/TNSRE.2019.2939655
- Yu, H., Li, F., Liu, J., Liu, D., Guo, H., Wang, J., et al. (2024). Evaluation of acupuncture efficacy in modulating brain activity with periodic-aperiodic EEG measurements. *IEEE Trans. Neural Syst. Rehabil. Eng.* 32, 2450–2459. doi:10.1109/TNSRE.2024.3421648
- Yu, H., Zeng, F., Liu, D., Wang, J., and Liu, J. (2025). Neural manifold decoder for acupuncture stimulations with representation learning: an acupuncture-brain interface. *IEEE J. Biomed. Health Inf.* 29 (6), 4147–4160. doi:10.1109/JBHI.2025.3530922
- Yuan, Y., Fu, R., Huang, L., Lin, W., Zhang, C., Chen, X., et al. (2021a). HRFormer: High-resolution vision transformer for dense prediction. *Adv. Neural Inf. Process. Syst.* 34, 7281–7293. doi:10.48550/arXiv.2110.09408
- Yuan, L., Chen, Y., Wang, T., Yu, W., Shi, H., Jiang, Z., et al. (2021b). “Tokens-to-token ViT: training vision transformers from scratch on ImageNet,” in *Proceedings of the IEEE international conference on computer vision* (New York, NY: ICCV), 558–567. doi:10.48550/arXiv.2101.11986
- Yuan, Z., Shao, P., Li, J., Wang, Y., Zhu, Z., Qiu, W., et al. (2024). YOLOv8-ACU: improved YOLOv8-pose for facial acupoint detection. *Front. Neurobotics* 18, 1355857. doi:10.3389/fnbot.2024.1355857
- Zhang, F., An, Q., Song, W., and Bai, M. (2024). Research on human acupoint detection by integrating key point information and acupoint theory. *IEEE Access* 12, 181889–181898. doi:10.1109/ACCESS.2024.3509026
- Zhou, E., Shen, Q., and Hou, Y. (2024). Integrating artificial intelligence into the modernization of Traditional Chinese Medicine industry: a review. *Front. Pharmacol.* 15, 1181183. doi:10.3389/fphar.2024.1181183
- Zong, D., Ding, C., Li, B., Li, J., Zheng, K., and Zhou, Q. (2023). “AcFormer: an aligned and compact transformer for multimodal sentiment analysis,” in *Proceedings of the ACM international conference on multimedia*, 2696–2704. doi:10.1145/3581783.3611974