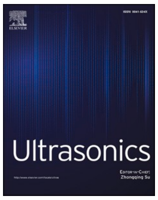
Ultrasonics 142 (2024) 107358

Contents lists available at ScienceDirect

Ultrasonics

journal homepage: www.elsevier.com/locate/ultras



Shear wave trajectory detection in ultra-fast M-mode images for liver fibrosis assessment: A deep learning-based line detection approach

Xinyi Wang [a], Bo Liu [d], Chonglin Wu [a], Zihao Huang [a], Yongjin Zhou [c], Xiaoming Wu [d], Yongping Zheng [a,b,*]

[a] Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region
[b] Research Institute for Smart Ageing, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region
[c] School of Biomedical Engineering, University of Shenzhen, Shenzhen, China
[d] Department of Computing, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region

ARTICLE INFO

Keywords:
Liver fibrosis
Transient ultrasound elastography
Line detection
Transformer
Deep learning

ABSTRACT

Stiffness measurement using shear wave propagation velocity has been the most common non-invasive method for liver fibrosis assessment. The velocity is captured through a trace recorded by transient ultrasonographic elastography, with the slope indicating the velocity of the wave. However, due to various factors such as noise and shear wave attenuation, detecting shear wave trajectory on wave propagation maps is a challenging task. In this work, we made the first attempt to use deep learning methods for shear wave trajectory detection on wave propagation maps. Specifically, we adopted five deep learning models in this task and evaluated them by using a well-acknowledged metric based on EA-Angular-Score (EAA) and task-specific metric based on Young s-Score (Ys) in the line-detection field. Furthermore, we proposed an end-to-end framework based on a Transformer and Hough transform, named Transformer-enhanced Hough Transform (TEHT). It took a wave propagation map as input image and directly output the slope of the shear wave trajectory. The framework extracts multi-scale local features from wave propagation maps, employs a deformable attention mechanism for feature fusion, identifies the target line using the Hough transform's voting mechanism, and calculates the contribution of each scale through channel attention. Wave propagation maps from 68 patients were utilized in this study, with manual annotation performed by a rater who was trained as a radiologist, serving as the reference value. The evaluation revealed that the SLNet model exhibited F-measure of EA and Ys values as 40.33 % and 40.72 %, respectively, while the TEHT model showed F-measure of EA and Ys values as 80.96 % and 98.00 %, respectively. TEHT yielded significantly better performance than other deep learning models. Moreover, TEHT demonstrated strong concordance with the gold standard, yielding R^2 values of 0.967 and 0.968 for velocity and liver stiffness, respectively. The present study therefore suggests the application of the TEHT model for assessing liver fibrosis owing to its superiority among the five deep learning models.

1. Introduction

Importance and difficulty of shear wave trajectory detection. Fibrosis is a disease that causes progressive stiffening of the liver, compromising its function and leading to over 1.2 million deaths annually [1]. Transient ultrasound elastography is a well-established non-invasive diagnostic imaging technique that uses shear waves to measure the stiffness of liver tissue for liver fibrosis assessment [2]. It generates wave propagation maps [3] by using ultra-fast A-mode ultrasound imaging (over 2000 frames per second) to form M-mode images as illustrated in Fig. 1. The stiffness of the liver is related to the shear wave velocity (SWV), which can be converted to Young's modulus [4], with the assumption of a linear, isotropic, incompressible material.

In previous studies, researchers have attempted to estimate the elastic value of liver by using various methods to calculate the SWV. Sandrin et al. utilized linear regression to analyze the evolution of signal delay [2]. The time-of-flight (TOF) technique has been validated and applied in several studies for tracking shear wave displacement data [2,5,6]. However, TOF methods are not robust in the presence of severe outliers, and the RANSAC algorithm has been proposed to improve

---

* Corresponding author at: Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region.
E-mail address: ypzheng@ieee.org (Y. Zheng).

https://doi.org/10.1016/j.ultras.2024.107358
Received 15 August 2023; Received in revised form 30 April 2024; Accepted 27 May 2024

detection robustness [7]. Rouze et al. introduced a method utilizing Radon Sum Transformation to SWV and demonstrated its real-time capabilities [8]. On the other hand, using the shear wave elasticity imaging (SWEI) technique, Li et al. proposed real-time tracking of trajectory curves to avoid issues about estimating shear wave velocity influenced by local maxima [9]. Meanwhile, Neidhardt et al. employed shear wave elastography (SWE) that bypassed shear wave velocity estimation and directly predicted local elasticity using 3D spatio-temporal CNNs from ultrasound data [10]. Not only methodological improvements have been made, but Song et al. also considered the issue from the perspective of ultrasound noise and improved motion detection of shear waves through pulse-inversion harmonic imaging [11]. Furthermore, Chen et al. innovatively proposed the use of MRCNN to denoise and enhance estimates from low-quality data [12], with the aim of improving the accuracy of estimating material properties such as elasticity or viscosity by increasing the signal-to-noise ratio (SNR) of SW particle motion. These efforts have made significant contributions to the detection of liver stiffness. However, due to various factors such as unclear imaging, noise, shear wave attenuation, and ultrasound attenuation, as illustrated in Fig. 1, detecting shear wave velocity by transient ultrasound elastography is still a challenging task, and current methods are not capable of effectively addressing these particular scenarios.

In Fig. 1, the first B-mode image and corresponding wave propagation map are chosen as reference due to the clear image quality without vascular interference in the B-mode image, and the presence of clean trajectories in the wave propagation map. In comparison, the second B-mode image, acquired from a patient with fatty liver disease, exhibits significant blurriness. For this commonly observed patient group, the corresponding wave propagation map demonstrates an unclear and fluctuating trajectory. In the third B-mode image, blood vessels can be easily observed in the region of interest. Noise interference shows in the wave propagation map when the shear wave passes blood vessels, making it challenging to accurately locate the shear wave trajectory. The fourth wave propagation map shows an irregular pattern (indicated by the white box), occurring when the shear wave propagation is disturbed by objects in the liver tissue or for some other reasons. More importantly, the presence of blurred texture on the shear wave trajectory in every wave propagation map (indicated by red box) can have a detrimental effect on the accuracy of shear wave elastography. The rightmost images in Fig. 1 represent an enlarged version of the region of interest for measurement (upper row) and an enlarged view of the blurred texture (lower row).

Shear wave trajectory detection via line detection on wave propagation maps based on deep learning: a new task in the medical domain. The slope of the shear wave trajectory in a wave propagation map is used to calculate SWV. Therefore, a new task that involves identifying the line and computing its slope on the wave propagation map based on deep learning for liver fibrosis assessment can be defined, which has not been explored before. Although there is no prior work that can be regarded as a reference for this task, we can utilize similar deep learning models designed for line detection on natural images in the other domains. The results of these models can be compared after training with our dataset, like Semantic Line Network (SLNet) [13], Line Convolution Neural Network (LCNN) [14] and Deep Hough Transform (DHT) [15]. SLNet employs contextual information and pairwise comparison to select lines, while LCNN extracts features through convolution and generates a junction heatmap, which is then utilized for selecting suitable lines. DHT utilizes the Hough transform and employs intersection point calculation to enhance the precision of line detection. However, these methods are limited in their ability to acquire long-distance feature dependency, which is crucial for detecting straight lines with irregular pattern. Moreover, heavy object detection methods, which are based on DEtection TRansformer (DETR) [16] like LinE segment TRansformers (LETR) [17] ignore the geometric properties of lines and may lead to unsatisfactory results and low efficiency.

Our proposed deep learning framework. In this paper, we design a domain-specific line detection network for the proposed task, named **T**ransformer-**e**nhanced **H**ough **T**ransform (TEHT). Specifically, wave propagation maps exhibit intricate and significant textures that play a crucial role in line detection. We refer to this characteristic as the “fine-grained property”. To accurately capture and retain this particular property within the input image, we commence by conducting a multi-scale feature extraction process. This process is inspired by a widely recognized framework called FPN [18], which allows us to retain features at multiple scales and thereby maximize the preservation of the aforementioned property. Within the extracted multi-scale features, the lower-scale features contain finer local texture information, while the higher-scale features capture broader global information in a more abstract manner. To better utilize these multi-scale features as well as learn a long-distance dependence, we propose to use the deformable-attention mechanism [19] for exploring semantic relationships in an efficient manner. Then, we convert attention-augmented convolutional space into a parameter space at each scale where we detect shear wave lines, similar to [15]. Finally, by fusing multi-scaled representations output from parameter space via channel attention, the line can be predicted and the corresponding slope can be calculated. To evaluate the effectiveness of the proposed framework and each module, we conducted comprehensive experiments and detailed ablation studies. The

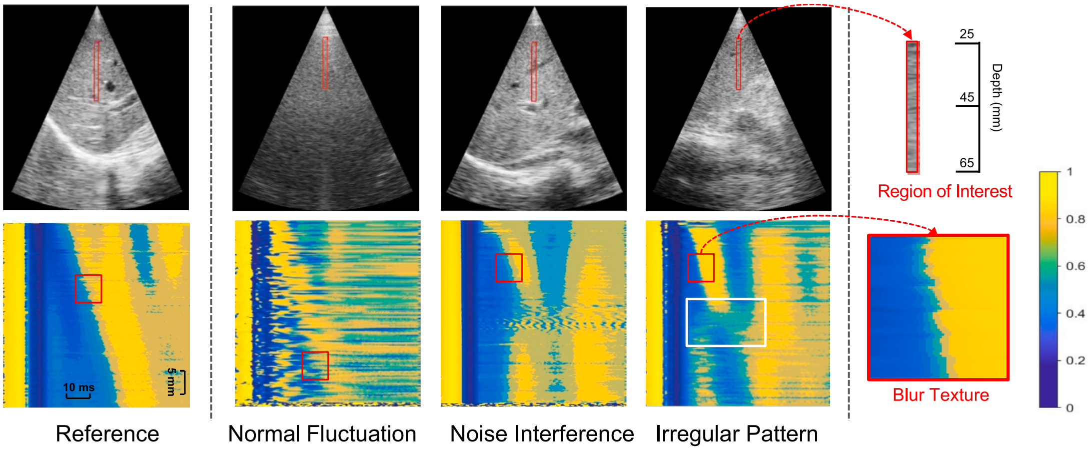


Fig. 1. Typical transient ultrasound wave propagation maps with different types of distortion in the trajectory of shear wave propagation. The corresponding B-mode image is displayed on the first row. The red rectangular box is the region of interest, which is the detection range with a depth of 25 mm–65 mm. The red box marks the area of the blur texture, which is observed in each map. The white box highlights a pattern that may be attributed to the encounter of wave propagation with objects. **(*It is recommended that this figure should be printed using the color.*)**

significant improvements over state-of-the-art methods in the general domain demonstrate the necessity to design a task-oriented method.

Our contributions are summarized as follows. (1) We successfully developed a deep learning-based approach to detect wave propagation traces on wave propagation maps for liver fibrosis assessment via line detection; (2) We proposed a domain-specific framework for this task as well as a tailored evaluation metric by taking the characteristics of a wave propagation map into consideration; (3) We conducted comprehensive experiments and achieved better results compared to state-of-the-art methods in the general domain.

2. Methods

Our proposed framework is motivated by the distinct properties of transient ultrasound wave propagation maps, which make them vastly different from natural images and thus represent a novel use of line detection in the medical domain. The framework includes a multi-scale wave feature extractor to represent fine-grained input images, a Hough transform module to learn a parametric domain for line detection, and a channel attention module for feature fusion, all of which are trained in an end-to-end manner for optimal performance.

2.1. Multi-scale wave feature extractor with deformable attention

Given an input wave propagation map $\mathbf{I} \in \mathbb{R}^{H\times W}$, we use ResNet-50 to extract multi-scale features $\mathbf{I}_{i\in\{1,2,3,4\}} \in \mathbb{R}^{H_i\times W_i\times C}$ with $H_i = \frac{H}{2^{i+1-\left\lfloor \frac{i}{4} \right\rfloor}}$ and $W_i = \frac{W}{2^{i+1-\left\lfloor \frac{i}{4} \right\rfloor}}$. To better learn long-range dependence between patches as well as utilize the knowledge at different scales, we propose to model relationships via a self-attention mechanism [20]. First, we use additional trainable 1×1 convolution layers to fix the number of channels at each scale to C'. Then, the features are further flattened into 2-D patches along the shape dimension and concatenated to $\in \mathbb{R}^{\sum_{i=1}^{4}(H_i\times W_i)\times C'}$. Then, we opt to use more efficient deformable attention [19], which could generate information-related sparse features $\widetilde{\mathbf{L}}$ to form keys $\widetilde{\mathbf{K}}$ and values $\widetilde{\mathbf{V}}$, where $\widetilde{\mathbf{L}}$ is extracted from the original $\mathbf{L}$ by using the learned position deviation (more details can be found in reference [19]). With complete feature $\mathbf{L}$ and sparse one $\widetilde{\mathbf{L}}$, the multi-head self-attention (MSA) layer can be formulated as follows:

$$\begin{cases} \mathrm{MSA_j}(\mathbf{L}) = \mathrm{Concat}(SA_1(\mathbf{L}), \dots, SA_h(\mathbf{L}))\mathbf{W^O}, \\ SA_i(\mathbf{L}) = \mathrm{Softmax}\left(\mathbf{Q_i}\mathbf{K_i} \middle/ \sqrt{D}\right)\mathbf{V_i}, \\ \mathbf{Q_i} = \mathbf{L}\mathbf{W_i^Q}, \mathbf{K_i} = \widetilde{\mathbf{L}}\mathbf{W_i^K}, \mathbf{V_i} = \widetilde{\mathbf{L}}\mathbf{W_i^V}, \end{cases} \tag{1}$$

where $\mathbf{W_i^Q} \in \mathbb{R}^{C'\times D}$, $\mathbf{W_i^K} \in \mathbb{R}^{C'\times D}$, and $\mathbf{W_i^V} \in \mathbb{R}^{C'\times D}$ are learnable weights, $1 \le i \le h$ where h is the number of heads, $1 \le j \le n$ where n is the number of MSA layers, and $\mathbf{W^O}$ is a learnable matrix used for final outputting.

After processing with MSA layers, we generate the final scale-cross representation $\widetilde{\mathbf{I}}$ with the same shape as $\mathbf{L}$, which is then decomposed and unflattened into four multi-scale feature maps with the same shape as that of the originals, i.e., $\widetilde{\mathbf{I}}_{i\in\{1,2,3,4\}} \in \mathbb{R}^{H_i\times W_i\times C'}$.

2.2. Hough transform for line detection

The Hough transform represents a line l in input space $\mathscr{X}(H,W)$ by a distance r to the center point (i.e., $r \in \left[-\sqrt{W^2+H^2}/2, \sqrt{W^2+H^2}/2\right]$) and an angle θ with the x-axis (i.e., $\theta \in [0,\pi)$). Given proper quantization intervals Δr and $\Delta\theta$ for r and θ, the line can be mapped to a point in a parameter space $\mathscr{P}(R,\Theta)$ with height $R = \frac{\sqrt{W^2+H^2}}{\Delta r}$ and width $\Theta = \frac{\pi}{\Delta\theta}$ via:

$$l' = (r/\Delta r, \theta/\Delta\theta) \tag{2}$$

Inspired by [15], we treat features $\widetilde{\mathbf{I}}_{\{1,2,3,4\}}$ as multi-scale input space and search potential lines in generated parameter spaces at each level $\widetilde{\mathbf{P}}_{i\in\{1,2,3,4\}} \in \mathbb{R}^{R_i\times\Theta\times C}$. Through comprehensive experiments, we set $R_i = \frac{100}{2^{i-1-\left\lfloor \frac{i}{4} \right\rfloor}}$ and $\Theta = 100$, respectively.

2.3. Feature aggregation via channel attention

The output $\widetilde{\mathbf{P}}_{\mathbf{i}}$ from the Hough transform module will first be upsampled to the same size and concatenated as $\mathbf{O}_{\mathrm{in}} \in \mathbb{R}^{R_1\times\Theta\times 4C}$. To reasonably amplify the contribution of information at each scale to the

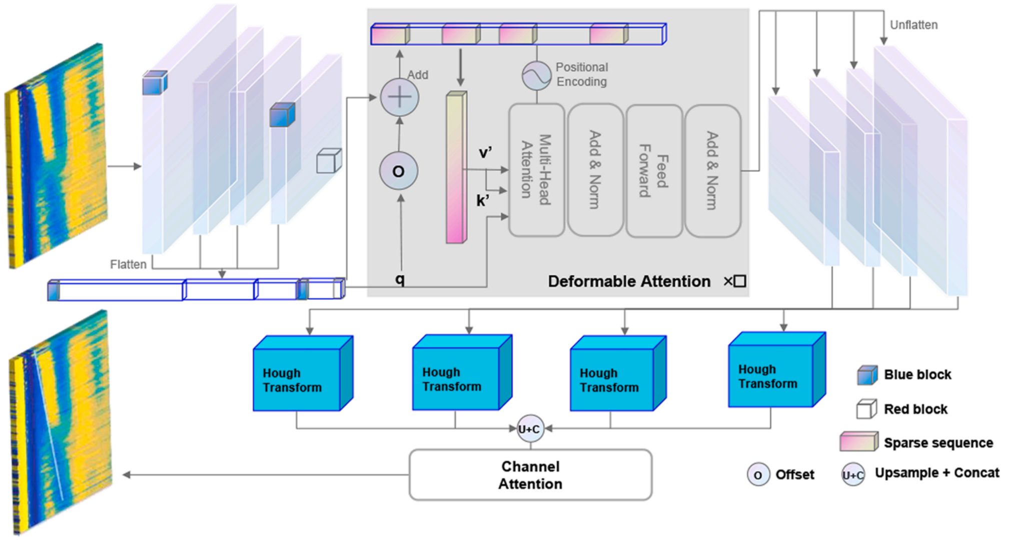


Fig. 2. Our proposed framework for shear wave trajectory detection on ultra-fast M−mode ultrasound images. The blue box and red box indicate feature pixels in different layers. The red block can be related to the blue block in the same or different layers, allowing the consideration of global information for model decision-making. ***(It is recommended that this figure should be printed using the color.)***

results, we weight each channel by channel attention [21]. In particular, the spatial information is compressed by parallel AvgPool and MaxPool as $\mathbf{M} \in \mathbb{R}^{1\times1\times4C}$. Then the channel attention map weighted output $\mathrm{CAM}_{\mathrm{output}}(\mathbf{O}_{\mathrm{in}})$ can be computed through a shared multi-layer perceptron (MLP) as follows:

$$\mathrm{CAM}_{\mathrm{output}}(\mathbf{O}_{\mathrm{in}}) = \sigma(\mathrm{MLP}(\mathrm{AvgPool}(\mathbf{O}_{\mathrm{in}})) + \mathrm{MLP}(\mathrm{MaxPool}(\mathbf{O}_{\mathrm{in}})))\mathbf{O}_{\mathrm{in}}, \tag{3}$$

where σ is the sigmoid activation layer.

Finally, $\mathrm{CAM}_{\mathrm{output}}(\mathbf{O}_{\mathrm{in}})$ is further processed through a 1×1 convolution layer to obtain the final result map $\mathbf{O}_{\mathrm{out}} \in \mathbb{R}^{R\times\Theta\times1}$, which will be used for computing the final pixel-wise cross-entropy loss:

$$Loss = -\sum \left(G\cdot log\sigma(\mathbf{O}_{\mathrm{out}}) + (1 - \mathbf{G})\cdot log(1 - \sigma(\mathbf{O}_{\mathrm{out}}))\right), \tag{4}$$

where $\mathbf{G} \in \mathbb{R}^{R\times\Theta\times1}$ is the hough map calculated from two point coordinates of the target line.

2.4. Evaluation metric

Our task is to calculate the liver stiffness using Young's modulus E. With the assumption of a linearly elastic, isotropic, homogeneous incompressible material with density ρ, the Young's modulus E is given by:

$$E = 3\rho V^2, \tag{5}$$

where ρ is the mass density set as 1000 kg/m^3 throughout the present study [2] and V is the shear wave velocity with the unit of meter per second (m/s). We estimate the shear wave velocity by first drawing a line in the wave propagation map; the slope of the line (represented by k) is numerically equivalent to the shear wave velocity. Then, we define θ as the angle between the line and x axis, ranging from 0° to 90°, and apply the following formula of $\tan(\theta)$ to estimate the slope of the line and hence its numerical equivalence of shear wave velocity.

$$k = tan(\theta), \tag{6}$$

The existing line detection evaluation metrics based on the distance between lines [14,22] are not appropriate for our task. A highly related metric to our task is EA-Score proposed in DHT [15]. They termed EA-Score, which considers both the Euclidean distance and Angular distance between a pair of lines. The Angular distance $\mathscr{S}_\theta$ is defined according to the angle between two lines (l_i, l_j):

$$\mathscr{S}_\theta = 1 - \frac{\theta(l_i, l_j)}{\pi/2}, \tag{7}$$

where $\theta(l_i, l_j)$ is the angle between l_i and l_j. The Euclidean distance is defined as:

$$\mathscr{S}_d = 1 - D(l_i, l_j), \tag{8}$$

where $\mathrm{D}(l_i, l_j)$ is calculated after normalizing the image into a unit square, representing the Euclidean distance between midpoints of l_i and l_j. In the end, the EA-Score is:

$$S = (\mathscr{S}_\theta \bullet \mathscr{S}_d)^2, \tag{9}$$

Based on this EA-Score, we can measure the matching level of any two lines (l_i, l_j). The higher the value of the EA-Score, the better the fitting performance between the two lines. However, the Young's modulus E only needs the shear wave velocity, which is related to Angular distance. Furthermore, if we consider that $\mathscr{S}_d$ should be higher, there is a possibility that the model may have a higher $\mathscr{S}$ but a lower $\mathscr{S}_\theta$, which would consequently lead to poorer ability in predicting E. In such case, the model may only be able to approximate the midpoint of the predicted line to the midpoint of the manually annotated line, rather than accurately predict the precise slope. Fig. 3**(a) and (b)** is an example showing the $\mathscr{S}_\theta$ and $\mathscr{S}_d$ of these two lines. Two predicted results for the same original image are shown on the left side of (a) and (b), with corresponding visual aid on the right side. It is evident that the predicted line in (a) is closer to the reference line and the angle between the two lines is smaller. Therefore, the $\mathscr{S}_\theta$ of (a) is higher than (b). However, due to the smaller distance between the midpoints of the two lines in (b), it has a higher $\mathscr{S}_d$ score, resulting in a higher $\mathscr{S}$. This is not consistent with our task expectations, so the equation needs to be modified to avoid the interference of $\mathscr{S}_d$. We set $\mathscr{S}_d$ as a constant value of 1 and modify the equation for EA-Score as follows, naming it EA-Angular-Score $\mathscr{S}_A$ to differentiate it from the previous equation:

$$\mathscr{S}_A = (\mathscr{S}_\theta \bullet 1)^2, \tag{10}$$

As the prediction performance improves, $\mathscr{S}_A$ increases from 0 to 1.

However, only using $\mathscr{S}_A$ to assess the performance of calculating Young's modulus can result in significant errors in some cases. Due to the nonlinear characteristics of $\tan(\theta)$, the same difference in θ can cause much larger errors in $\tan(\theta)$ when θ falls into the range $(\pi/3, 2\pi/3)$ compared to outside of it. Nevertheless, the $\mathscr{S}_A$ only evaluates the model's performance based on the angle difference $\theta(l_i, l_j)$, which has the possibility to select the model with more angle difference while θ is in the range $(\pi/3, 2\pi/3)$, resulting in significant errors in calculating E. Furthermore, the high value of Young's modulus brought by θ in this range indicates severe liver fibrosis, meaning that high precision is demanded and such errors are unacceptable. Therefore, we directly observe $\tan(\theta)$ and propose a tailored Young's-Score $\mathscr{S}_Y$ as follows:

$$\mathscr{S}_Y = \left(\tan(\theta_{\mathrm{reference}}) - \tan(\theta_{\mathrm{output}})\right)^2, \tag{11}$$

where $\tan(\theta_{\mathrm{reference}})$ and $\tan(\theta_{\mathrm{output}})$ are the slopes of a pair of lines to be measured. To avoid the situation that the angle distance is large but Young's modulus are the same, we choose $\tan(\theta)$ instead of $\tan(\theta)^2$. $\mathscr{S}_Y$ increases from 0 when the detection accuracy decreases.

In Fig. 4, different prediction results for the same image are shown, demonstrating that as the prediction results improve, the Young's-Score decreases and the EA-Angular-Score increases.

2.5. Statistical analyses of slope and liver stiffness

To evaluate the accuracy of TEHT in measuring slope and liver stiffness, we conducted Bland-Altman analysis [23] and Pearson correlation [24].

3. Experiments

3.1. Experimental setup

3.1.1. Datasets

The dataset used in this study consists of 1,751 wave propagation maps collected from 68 patients with varying degrees of chronic liver diseases. Because 8 patients withdrew from the consent, we ended up using data from only 60 patients. The 3.5-MHz phased array-based transient elastography probe was used for wave propagation map data acquisition under the guidance of real-time B-mode ultrasound imaging, using the in-house developed transient elastography system[3]. The elastic wave acquisition frame rate is 6400 frames/second. The vertical axis of the sample is the depth with 25–65 mm, while the horizontal axis is the time range from 0 to 80 ms. As per the procedure outlined in prior studies, the participants were instructed to assume a dorsal decubitus position, with their right arm fully extended and positioned behind their heads. The transducer tip of the probe was then positioned on the skin surface between the ribs, aligning with the right lobe of the liver. Utilizing real-time B-mode imaging guidance, an optimal acoustic window

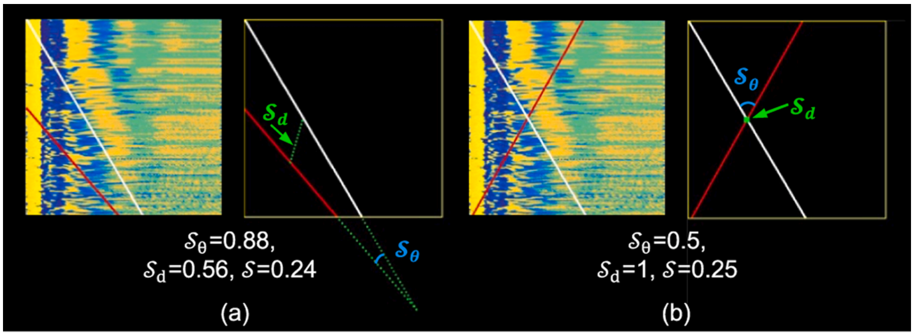


Fig. 3. Examples for explaining the EA-Score. The white line represents the reference line manually marked by a rater who was trained as a radiologist with five years of experience, while the red line represents the predicted line. The green dashed lines are auxiliary lines to facilitate an easier understanding of the image, including the midpoint connection line and the extension line. The values of $\mathscr{S}_\theta$ and $\mathscr{S}_d$ are calculated based on the information provided in the images and formulas 8 and 9. **(It is recommended that this figure should be printed using the color.)**

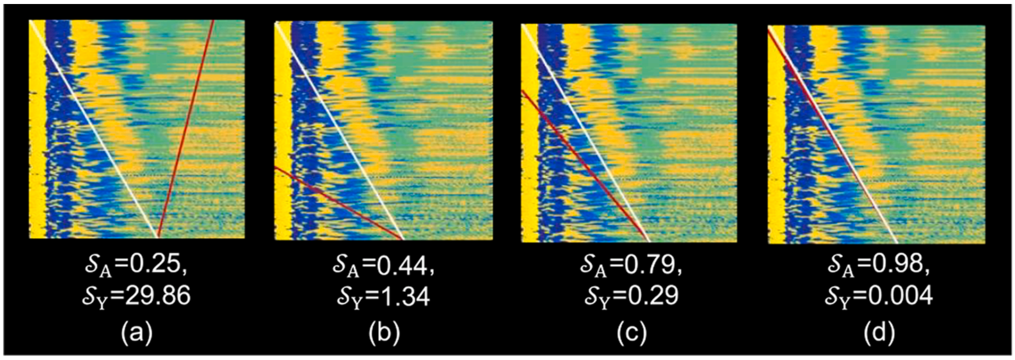


Fig. 4. Examples for explaining the EA-Angular-Score and Young's-Score. **(It is recommended that this figure should be printed using the color.)**

that provided clear visualization of the liver parenchymal region was selected. Because the liver constantly moves due to respiration, the wave propagation maps obtained from the same subject can be very different. To exclude possible interference, 25–30 measurements for one patient was taken on the same day to make sure at least 10 valid images. We split these data on subject level into a training set and a test set with a ratio of 60 %/40 %. The shear wave trajectory in wave propagation maps was identified and annotated by a rater who was trained as a radiologist and has 5 years of experience in liver ultrasound elastography, using version 4.5.6 of the Labelme software [25]. The slope of shear wave trajectory manually drawn was then converted to SWV which served as the reference value in this study.

Ethical approval was granted by the Institutional Review Board (IRB) of The Hong Kong Polytechnic University (HSEARS20210809002). Informed consents were obtained from all the enrolled subjects.

3.1.2. *Subject demographics*

During the period between May and October 2022, a total of 68 consecutive adult patients with known chronic liver diseases were recruited for this study. Of these, the characteristics of 68 subjects were considered for the final statistical analyses, and their summarized information can be found in Table 1. The overall median age of the participants was 54 years, with a range of 22 to 73 years (IQR of 18). Among the subjects, 51 % were male. The median body mass index (BMI) was 23 kg/m^2, with an IQR of 5 and a range of 16–39. It is worth noting that nearly 30 % of the patients fell into the overweight or obese category, as defined by a BMI of 25 kg/m^2 or higher. The median waist circumference was measured to be 86 cm. The majority of the participants (73 %) were diagnosed with chronic hepatitis B, with 32 % of them also experiencing coexistent steatosis. Additionally, 27 % of the subjects were found to have non-alcoholic fatty liver disease (NAFLD).

Table 1
Characteristics of the study cohort (n = 68).

Characteristics	Median [IQR] or Proportion % (n)
• Demographics	
Female	49 % (33)
Age, years	54 [18; range 22–73]
Overweight & obese	29 % (20)
Central obesity	52 % (35)
Metabolic syndrome	7 % (10)
• Anthropometrics	
BMI, kg/m^2	23 [5; range 16–39]
Waist circumference, cm	86 [15; range 65–115]
Skin-to-liver capsule distance, mm	13 [6, range 5–25]
• Liver disease etiology	
HBV	48 % (33)
NAFLD	27 % (18)
Coexistence of HBV and NAFLD	25 % (17)

IQR = interquartile range; BMI = body mass index; NAFLD = non-alcoholic fatty liver disease; HBV = hepatitis B virus.

3.1.3. *Implementation details*

The Labelme software (Massachusetts Institute of Technology, Cambridge, MA, USA) is used to annotate data. Our model is trained on an RTX 3090 GPU (Nvidia Corporation, Santa Clara, CA, USA) for 100 epochs, with a batch size of 12. We utilize the Adam optimizer to minimize the loss function, with a momentum of 0.9, a learning rate of 2×10^{-4}, and no weight decay. The number of attention heads is 8 and the number of MSA layers is 6 in the proposed model. The input image size is 400 × 400, and a left–right flip is applied to augment the input data. To ensure consistent class distribution, we conduct five-fold cross-validation in the proposed model.

3.2. Results

3.2.1. Comparison with state-of-the-art models

In this section, we evaluate our model's performance and four cutting-edge line detection methods in the non-medical domain: SLNet [13], LCNN [14], LETR [17], and DHT [15]. We use precision, recall, and F-measure as evaluation metrics, based on the EA-Angular-Score proposed in the DHT[15] and our newly proposed Young's-Score. Deep learning models have shown great potential in detecting SWV. TEHT outperforms the other methods in terms of detecting the number of lines and angle precision, as demonstrated in Table 2. The evaluation revealed that the SLNet model exhibited F-measure of EA-Angular-Score and Young's-Score values as 40.33 % and 40.72 %, respectively, while the TEHT model showed an F-measure of EA and Ys values of 80.96 % and 98.00 % respectively. Compared with the second-best results from state-of-the-art models in other domains, our model improves these scores by 1.91 % and 1.87 %.

3.2.2. Ablation study for evaluating impact of different components

To assess the contribution of components just in the TEHT model, we conducted a validation process to evaluate the effectiveness of each individual element. Table 3 shows that the multi-scale wave feature extractor alone improves the F-measure by 4.21 % to 96.66 %. When only the channel attention module is used without the multi-scale wave feature extractor, the performance reaches 97.15 %. Combining both components achieves the best performance of 98 %, indicating this combination is the most recommended architecture. Please note that the Hough transform is the basic module which must be included.

3.2.3. Qualitative evaluation

We visualize the results predicted by our proposed TEHT along with LETR and DHT in Fig. 5 and select five representative wave propagation maps to showcase the varied image quality, which we mentioned in the introduction. Each row in this figure represents the output results of the same model applied to the wave propagation maps of five patients, while each column represents the output results of three different models applied to the same wave propagation map. The white line in the figure is manually annotated by a rater, who is trained as a radiologist and has five years of experience in liver ultrasound elastography. The slope of this annotated line serves as the reference value. The red line, on the other hand, represents the result automatically detected by the model. The closer the red line is to the white line, the better the model's performance.

For patients A and B, the images exhibit typical characteristics of obese subjects who are more prone to liver fibrosis. In column A, the DHT model fails to detect any lines, limiting its applicability for such patients. So, we can only find one white line in the middle row image without any red line. Both the LETR and TEHT models successfully detect a single red line, with the predicted line of TEHT almost perfectly aligning with the manually annotated line. In the case of patient B, both the DHT and TEHT models predict lines that closely match the manually annotated line, while the LETR model exhibits larger errors. Although this error may not be very noticeable visually, the difference between the reference stiffness and the predicted liver stiffness from LETR is 1.75 kPa, which exceeds the acceptable error range.

Table 2
Comparison of our TEHT and state-of-the-art methods for line detection. The best and second-best results are bolded and underlined, respectively.

	EA-Angular-Score (%)			Young's-Score (%)		
	Precision	Recall	F-measure	Precision	Recall	F-measure
SLNet	39.83	40.85	40.33	40.21	41.25	40.72
LCNN	59.43	64.95	62.07	70.85	77.43	74.00
LETR	73.42	73.42	73.42	96.13	96.13	96.13
DHT	79.26	78.85	79.05	95.98	95.49	95.73
TEHT	**80.96**	**80.96**	**80.96**	**98.00**	**98.00**	**98.00**

Table 3
Effectiveness of different components in our TEHT framework. √ denotes the usage of specific components in the model.

Hough transform	Multi-scale wave feature extractor	Channel attention	F-measure (%)
√			92.45
√	√		96.66
√		√	97.15
√	√	√	98.00

For patients C and D, the images exhibit noise interference, which was mentioned in the introduction. Both the DHT and LETR models are affected by this interference, leading to inaccurate detection. In contrast, the proposed TEHT consistently demonstrates the best detection results, showcasing excellent resistance to interference. In (e), the irregular pattern mentioned in the introduction becomes evident. According to our test experiment on the DHT model, we found that this model struggled to select the most appropriate line in such scenarios, resulting in multiple output results. Therefore, two red lines are present in the middle row image, representing two prediction outcomes. This multiple output result can be confusing for users. However, both the LETR and TEHT models accurately predict reasonable lines.

Faced with varied image quality and complex situations, TEHT consistently exhibits the highest accuracy.

3.2.4. Statistical results

As shown in Fig. 6**(a-d)**, our method achieves high consistency with manual measurements. The average bias and correlation coefficient (R^2) for line slope are −0.003 and 0.967, respectively, while those for liver stiffness are −0.023 and 0.968. These results demonstrate the superiority of our TEHT and its potential for measuring liver stiffness.

We further conducted analysis of the DHT model for detecting slopes and liver stiffness results. The DHT model was unable to detect any lines in cases with poor image quality, which accounted for approximately 5 % of the cases. Fig. 6**(e-h)** presents the analysis results of the DHT model after excluding these unpredictable cases. In (e) and (g), it can be observed that the average bias of difference between DHT automated measurements and reference value are slightly lower or equal to the average bias displayed in (a) and (c). However, the consistency range is approximately twice as wide as the consistency range of TEHT results and the reference value. Fig. 6**(f) and (h)** demonstrate R^2 of 0.896 for line slope and 0.889 for liver stiffness, respectively.

LETR exhibits similar capabilities to TEHT in predicting results in all cases. Fig. 6**(i-l)** presents the analysis of LETR based on 100 % of cases. The average bias and R^2 for line slope are 0.035 and 0.811, respectively, while those for liver stiffness are 0.284 and 0.816. In (i) and (k), it can be observed that the average bias of difference between LETR automated measurements and reference value is significantly higher compared to the average bias displayed in (a) and (c). Moreover, the R^2 for line slope and liver stiffness, as shown in (j) and (l) respectively, are not sufficiently high to be widely accepted for practical use.

4. Discussion

Extensive research has confirmed the value of using shear wave propagation velocity for non-invasive detection of liver fibrosis [2,8,11,26], and significant contributions have been made from various directions in this field. These studies have also highlighted the influences of factors such as imaging noise, shear wave attenuation, and ultrasound attenuation on detecting shear wave motion, thus posing significant challenges in line detection tasks on SWV. Deep learning models are capable of identifying abstract or high-level features (ref).

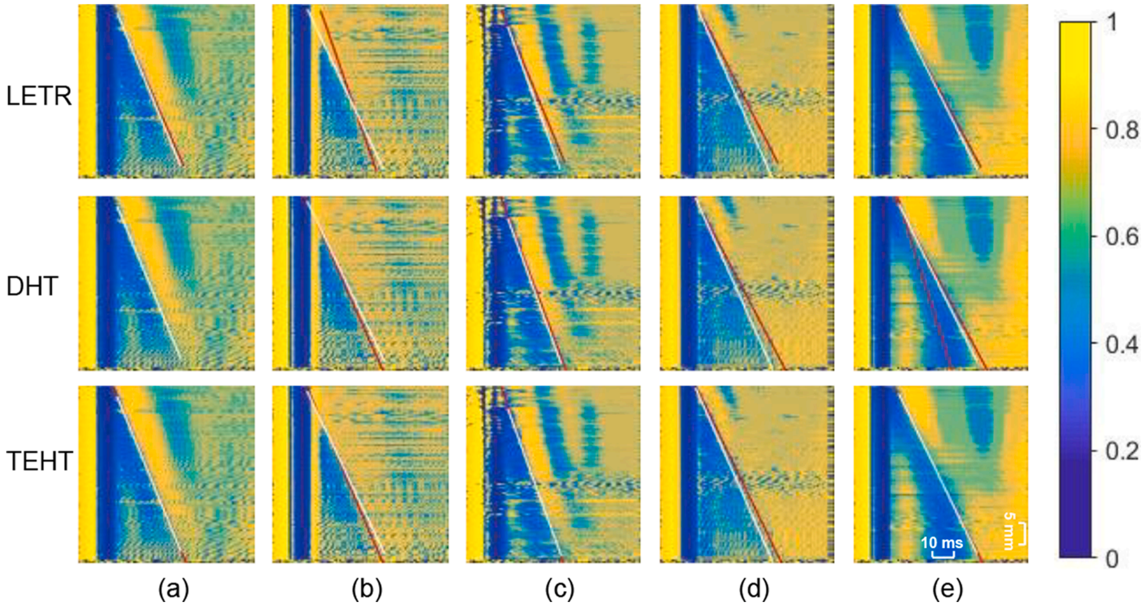


Fig. 5. Visualization of the results on five representative wave propagation maps, with the expert's manually annotated white line and the predicted line represented by red line. ***(It is recommended that this figure should be printed using the color.)***

Therefore, our research introduced deep learning models to this field with the aim of improving the detection of SWV. Based on our research, we found that our proposed deep learning model performed well in regard to detecting SWV under various conditions shown in Fig. 1. In particular, the TEHT model achieved an R^2 of 0.96 compared to the reference, as shown in Fig. 6. This finding demonstrated the effectiveness of the THEH model in automatic SWV detection, which would greatly assist the assessment of liver fibrosis.

4.1. Superiorities of the proposed TEHT model over the existing methods studied

The comparative analysis in the present study demonstrated that the proposed TEHT model was the best-performing method in terms of the EA-Angular-Score and the Young's-Score, followed by the existing DHT [15] and LETR [17] methods, while the SLNet and LCNN models were the most under-performing methods (Table 2).

We speculated that the superiority of our proposed TEHT method over the DHT model could be ascribed to the capability of the TEHT in incorporating long-distance features, hence enabling a more effective learning of global information. For instance, Fig. 5 **(e)** shows that the DHT method detected two red lines on the wave propagation map with irregular patterns, where global information of the map may be more crucial in accurate line detection. A possible reason behind may partly be attributed to the limitation of the DHT model in learning and integrating global information on the shear wave propagation map. In this circumstance, the DHT model might mainly focus on the local features from the front and back segments of the shear wave propagation map. Indeed, the DHT model employs a conventional neural network for feature extraction, which is deficient in learning long-distance dependencies [20]. To overcome this challenge, we introduced a deformable attention mechanism in the feature-extracting encoder in our proposed TEHT model, for the sake of enhancing its capability in learning and incorporating long-distance features [19]. In terms of the network architecture of the proposed model **(**Fig. 2**)**, the deformable attention module enables a more effective complementary learning of the flattened features represented in different boxes as they are processed through the multi-head attention mechanism. This capacity would still hold in the case between the leftmost blue box and the rightmost red box in Fig. 2, which represents the farthest apart of the flattened features extracted by the multi-scale wave feature extractor. In other words, the introduced deformable attention module enhances the complementary learning between these blocks for capturing global information. Consequently, our proposed TEHT was capable of accurately detecting a single line with a good agreement with the reference line in the case of irregular wave propagation pattern, as shown in Fig. 5 **(e)**.

Apart from this, the superiority of the TEHT model over the existing LETR approach may in part be ascribed to the advantage of the Hough Transform module incorporated in the TEHT model for a more effective learning of line's geometric features. For example, Fig. 5 **(b)** shows that the LETR exhibits deficiencies in accurate line detection compared to the TEHT model in the case of unclear image. A possible explanation behind may be related to the fact that the existing LETR approach is based on the DETR, which was designed for predicting bounding boxes, rather than focusing on the geometric properties of a straight line [16,17], which may in turn lead to inferior line detection in unclear image. By contrast, our TEHT model enables a more effective learning of line's geometric features owing to the incorporation of the Hough Transform module which was dedicatedly designed for line detection. Therefore, the TEHT model yielded a superior detection accuracy compared to the LETR method in the case of unclear image, as displayed in Fig. 5 **(b)**.

Notably, the TEHT consistently exhibits the most accurate detection capabilities in the five representative images of Fig. 5, despite the presence of factors such as unclear imaging and noise interference. Our proposed model not only overcame the abovementioned limitations of the DHT and LETR models, but it also yielded tremendous improvements in line detection performance compared to the SLNet [13] and the LCNN [14] methods **(**Table 2**)**.

4.2. Data augmentation

Data augmentation techniques have been widely employed to improve the performance of benchmarked datasets and to increase the quantity and diversity of training data[27]. Common image augmentation methods are based on image transformations, such as flipping, rotating, and blurring[28,29,30,31,32]. In addition, noise injection involves adding noise to signals at different signal-to-noise ratios (SNRs), which accounts for the stochastic processes in electronic devices of ultrasound transducers and measurement systems [33]. Synthesis techniques enhance data diversity by generating artificial images, with

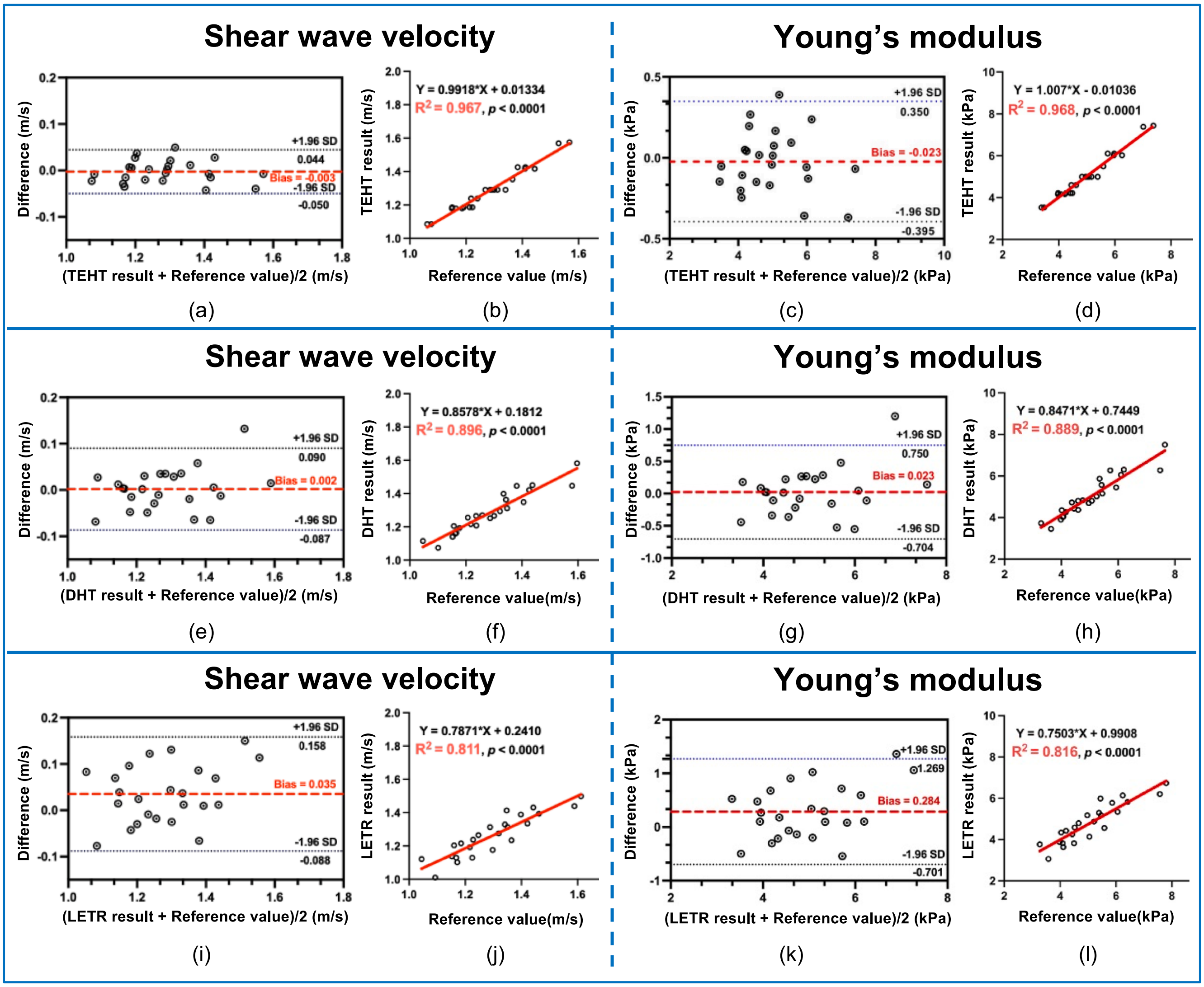


Fig. 6. Assessment of consistency between our TEHT and manual measurements in (a-d). Assessment of consistency between DHT and manual measurements in (e-h). The DHT model has problem to detect any line in 5% cases with poor image quality, resulting this figure is based on 95% images. Assessment of consistency between LETR and manual measurements in (i-l). Bland-Altman analysis is depicted in (a & c & e & g & i & k), while Pearson correlation is shown in (b & d & f & h & j & l). ***(It is recommended that this figure should be printed using the color.)***

methods like mixup [34] representing synthetic sample data augmentation [35,36,37,38,39].

However, there is limited research focused on data augmentation approaches based on shear wave propagation maps. In practice, operations like rotation and translation are beneficial as they generate images that occur in real-world scenarios, whereas flipping and cropping may have limited gains as the generated images are less common in practical operations. Noise injection as a data augmentation technique may yield significant gains, as noise poses a major challenge to model detection. The methods of generating noise are worth exploring, such as adding noise to signals, applying Gaussian noise to images, and creating noise patterns referred to as noise interference in the Fig. 1. Synthetic techniques will bring significant benefits if it can generate various scenarios mentioned in the Fig. 1 and produce more potential cases. Although they are not real acquired images, they attempt to mimic the texture of real images and greatly increase the samples corresponding to rare cases of liver stiffness. Lastly, evaluating the effectiveness of different augmentation techniques requires rigorous experiments. In addition to comparing the magnitude of results obtained from different augmentation techniques, the impact of data volume needs to be considered. Examples of questions include whether or not the order of magnitude of enhancement is the same for different augmentation techniques at 100 % and 50 % data volumes? Whether or not slight enhancement observed at 100 % data volume is due to the limitations of the augmentation techniques or because the data volume is already sufficiently diverse? Which augmentation technique would yield the greatest improvement on results at different data volumes? Answering these questions can assist researchers in selecting the most suitable augmentation techniques for their specific priority datasets.

4.3. Minimal improvement

Deep learning algorithms can extract abstract features and perform self-learning from manually labeled data, resulting in high accuracy in detecting lines for SWV. For example, DHT and LETR achieved Young's-Score of 95.73 % and 96.13 % on our test dataset, with correlation coefficients of 0.896 and 0.811, respectively, compared to the manually estimated shear wave velocity. These metrics indicate that the models achieved high accuracy in line detection, although not perfect. This could be attributed to the accumulated technical advancements in these models. However, there are still areas for improvement in DHT and LETR. For instance, they may fail to detect lines in low-quality images, limiting the examination of certain patients, or they may detect two lines in a single image, causing confusion for the users.

To address these limitations, we proposed the TEHT model for further optimization. Although the numerical improvement compared to the existing models was not substantial, enhancing accuracy from 96 % to 98 % is far more challenging than from 76 % to 96 %. This issue of small improvement margins also arises in the Table 3. We examined the roles of various components in our proposed model through this experiment. From Table 3, it can be observed that when the score of the base structure is 92.45 %, adding either the Multi-scale Wave Feature Extractor or the Channel Attention module results in at least 4 % improvement. However, when one module is added, the improvement of the other module becomes minimal. It remains to be determined whether these small growth rates were due to limitations in the model or module capabilities or because of the high base score, which can be further investigated through additional experiments. For example, when the dataset is rich in sources and abundant in quantity, line detection tasks will likely present more challenges and diversity. This provides a significant opportunity for improving the results. Therefore, by revalidating the existing methods on new datasets, we can clearly demonstrate the effectiveness of the methods through significant improvements in numerical representation. The potential significance of our model can also be quantified more explicitly. Although the dataset included in this study is limited on diversity, we demonstrated the feasibility of the proposed method for detecting shear wave trajectories to estimate shear wave velocity. Nevertheless, a larger cohort encompassing data from multiple medical centers and countries is warranted to validate our proposed methods further.

4.4. Inhomogeneous case

When evaluating liver stiffness, it is important not only to consider homogeneous liver cases but also to take into account inhomogeneous cases. [40] performed biopsies on the right and left lobes of the liver in patients, and more than 30 % of the patients showed discordant fibrosis stages. Similarly, [41] studied 51 cases of NAFLD and found variability in the results of two liver biopsies, with 41 % of patients showing different stages of fibrosis. Several similar studies [40,41,42,43] also demonstrated liver heterogeneity, resulting in different estimations of elasticity in different regions. Therefore, predicting liver stiffness by a suitable way in inhomogeneous cases is a challenge. One limitation of ARFI and TE techniques is that most of the methods based on these technologies are one-dimensional, making it difficult to measure variability [44]. Consequently, the typical approach to determine a representative liver stiffness value is to calculate the mean or median value [45]. On the other hand, SSI and SWE techniques can provide two-dimensional images that visualize the spatial heterogeneity of liver stiffness [46,47]. However, SWE has certain limitations in terms of imaging depth and spatial resolution. Additionally, validating the expression of liver heterogeneity is also challenging, as it requires biopsies from different parts of the liver, which is difficult to achieve in practice. Our technique is based on TE and overcomes the limitation of being one-dimensional by providing two-dimensional images, which enables the possibility of reflecting local viscoelasticity changes. However, this article represents our first step in using shear wave propagation maps to detect shear wave velocity, following the tradition [2,48,49,50,51] to use a straight line as a measurement. Nonetheless, it is highly recommended to explore appropriate methods to represent liver heterogeneity in a larger cohort containing sufficient inhomogeneous cases in future studies. For example, it is interesting and pertinent to explore deep learning-based segmentation or tracking methods to detect curved wavefronts, or to employ generative techniques for providing more detailed representations of the differences in elastic stiffness in various liver regions.

4.5. Limitations

However, several limitations existed in our research.

First of all, although this study enrolled a large sample size of 1,751 wave propagation maps to safeguard the statistical power of our inference from the proposed methods, it might still be limited by the fact that all the enrolled participants were recruited in Hong Kong, as subjects from different geographic regions and races might exhibit different patterns of wave propagation maps on top of their liver conditions [52]. Therefore, future studies involving a more diverse races from different geographic regions would further strengthen generalizability of our findings in a global scale, which will be part of our future study.

Secondly, although our proposed method demonstrated encouraging performance compared to other state-of-the-art models, its high computational cost (currently occupies 349.8 MB) may appear to be a shortcoming when our proposed method is implemented in portable devices, which are widely popular for clinical users during outdoor scanning. Its computational efficiency can be further enhanced to facilitate clinical adoption by utilizing distillation techniques [53] to lower memory usage.

Apart from these limitations, our study presents a significant potential for further exploration of using deep learning to detect SWV, as it can serve as a baseline. We conducted detailed experiments on existing models in the non-medical domain and proposed an efficient yet simple model that can be easier reproduced. Furthermore, we introduced an evaluation metric that takes into consideration the specific characteristics of this task. We believe that these works would serve as a valuable reference for other researchers, facilitating the development of new deep learning models based on elastography and promoting the integration of computer technology with ultrasound imaging techniques.

5. Conclusion

We introduced a novel approach to shear wave trajectory detection for fibrosis assessment via line detection based on deep learning. Our study presents the first attempt to apply deep learning methods to detect shear wave trajectories on wave propagation maps. Our proposed solution is an end-to-end framework that incorporates Transformer and the Hough transform. The framework extracts multi-scale local features from fine-grained wave propagation maps, applies a deformable attention mechanism for effective feature fusion, and uses channel attention to calculate the contribution of each scale. Experiments on an ultrasound wave propagation map dataset containing 1751 images with varying liver conditions demonstrated the effectiveness of our proposed approach, which would serve as a reference for future research.

CRediT authorship contribution statement

Xinyi Wang: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Bo Liu:** Writing – review & editing. **Chonglin Wu:** Validation. **Zihao Huang:** Formal analysis, Data curation. **Yongjin Zhou:** Conceptualization. **Xiaoming Wu:** Writing – review & editing, Supervision. **Yongping Zheng:** Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This study has been partially supported by the Research Institute for Smart Ageing of the Hong Kong Polytechnic University (1-CD5B).

References

[1] I.I. Abubakar, T. Tillmann, A. Banerjee, Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013, Lancet 385 (9963) (2015) 117–171.

[2] L. Sandrin, et al., Transient elastography: a new noninvasive method for assessment of hepatic fibrosis, Ultrasound Med. Biol. 29 (12) (2003) 1705–1713.

[3] T.-M. Mak, Y.-P. Huang, Y.-P. Zheng, Liver fibrosis assessment using transient elastography guided with real-time B-mode ultrasound imaging: a feasibility study, Ultrasound Med. Biol. 39 (6) (2013) 956–966.

[4] L. Euler, The rational mechanics of flexible or elastic bodies 1638-1788: introduction to Vol. X and XI. Springer Science & Business Media, 1980.

[5] M. Tanter, et al., Quantitative assessment of breast lesion viscoelasticity: initial clinical results using supersonic shear imaging, Ultrasound Med. Biol. 34 (9) (2008) 1373–1386.

[6] M.L. Palmeri, M.H. Wang, J.J. Dahl, K.D. Frinkley, K.R. Nightingale, Quantifying hepatic shear modulus in vivo using acoustic radiation force, Ultrasound Med. Biol. 34 (4) (2008) 546–558.

[7] M.H. Wang, M.L. Palmeri, V.M. Rotemberg, N.C. Rouze, K.R. Nightingale, Improving the robustness of time-of-flight based shear wave speed reconstruction methods using RANSAC in human liver in vivo, Ultrasound Med. Biol. 36 (5) (2010) 802–813, https://doi.org/10.1016/j.ultrasmedbio.2010.02.007.

[8] N.C. Rouze, M.H. Wang, M.L. Palmeri, K.R. Nightingale, Robust estimation of time-of-flight shear wave speed using a radon sum transformation, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 57 (12) (2010) 2662–2670.

[9] Y. Li, Q. Lv, J. Dai, Y. Tian, J. Guo, Shear wave velocity estimation using the real-time curve tracing method in ultrasound elastography, Appl. Sci. 11 (5) (2021) 2095.

[10] M. Neidhardt, et al., Ultrasound shear wave elasticity imaging with spatio-temporal deep learning, I.E.E.E. Trans. Biomed. Eng. 69 (11) (2022) 3356–3364, https://doi.org/10.1109/TBME.2022.3168566.

[11] P. Song, et al., Improved shear wave motion detection using pulse-inversion harmonic imaging with a phased array transducer, IEEE Trans. Med. Imaging 32 (12) (2013) 2299–2310.

[12] X. Chen, N. Chennakeshava, R. Wildeboer, M. Mischi, R.J.G. van Sloun, Shear-wave particle-velocity estimation and enhancement using a multi-resolution convolutional neural network, Ultrasound Med. Biol. 49 (7) (2023) 1518–1526, https://doi.org/10.1016/j.ultrasmedbio.2023.02.004.

[13] J.-T. Lee, H.-U. Kim, C. Lee, and C.-S. Kim, "Semantic line detection and its applications, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 3229–3237.

[14] Y. Zhou, H. Qi, Y. Ma, End-to-end wireframe parsing, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 962–971.

[15] K. Zhao, Q. Han, C.-B. Zhang, J. Xu, M.-M. Cheng, Deep hough transform for semantic line detection, IEEE Trans. Pattern Anal. Mach. Intell. 44 (9) (2021) 4793–4806.

[16] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko, End-to-end object detection with transformers, in: European conference on computer vision, Springer, 2020, pp. 213–229.

[17] Y. Xu, W. Xu, D. Cheung, Z. Tu, Line segment detection using transformers without edges, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 4257–4266.

[18] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, Feature pyramid networks for object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2117–2125.

[19] Z. Xia, X. Pan, S. Song, L.E. Li, G. Huang, Vision transformer with deformable attention, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 4794–4803.

[20] A. Vaswani, et al., Attention is all you need, Adv. Neural Inf. Process Syst. 30 (2017).

[21] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, Cbam: Convolutional block attention module, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 3–19.

[22] K. Huang, Y. Wang, Z. Zhou, T. Ding, S. Gao, Y. Ma, Learning to parse wireframes in images of man-made environments, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 626–635.

[23] D.G. Altman, J.M. Bland, Measurement in medicine: the analysis of method comparison studies, J. Royal Stat. Soc. Ser. D: Statistician 32 (3) (1983) 307–317.

[24] K. Pearson, VII. Note on regression and inheritance in the case of two parents, in: Proceedings of the royal society of London, vol. 58, no. 347–352, 1895, pp. 240–242.

[25] A. Torralba, B.C. Russell, J. Yuen, Labelme: Online image annotation and applications, Proc. IEEE 98 (8) (2010) 1467–1484.

[26] A. Ozturk, M.C. Olson, A.E. Samir, S.K. Venkatesh, Liver fibrosis assessment: MR and US elastography, Abdominal Radiology (2022) 1–14.

[27] Q. Wang, et al., Dynamic data augmentation based on imitating real scene for lane line detection, Remote Sensing 15 (5) (2023), https://doi.org/10.3390/rs15051212.

[28] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, J. Garcia-Rodriguez, A survey on deep learning techniques for image and video semantic segmentation, Appl. Soft Comput. 70 (2018) 41–65.

[29] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (5786) (1979, 2006,) 504–507.

[30] X. Yu, X. Wu, C. Luo, P. Ren, Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework, Gisci Remote Sens 54 (5) (2017) 741–758.

[31] A.S. Parihar, O.P. Verma, C. Khanna, Fuzzy-contextual contrast enhancement, IEEE Trans. Image Process. 26 (4) (2017) 1810–1819.

[32] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[33] S. Uhlig, I. Alkhasli, F. Schubert, C. Tschöpe, M. Wolff, A review of synthetic and augmented training data for machine learning in ultrasonic non-destructive evaluation, Ultrasonics 134 (2023) 107041, https://doi.org/10.1016/j.ultras.2023.107041.

[34] H. Zhang, M. Cisse, Y. N. Dauphin, D. Lopez-Paz, mixup: Beyond empirical risk minimization, arXiv preprint arXiv:1710.09412, 2017.

[35] D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, C.A. Raffel, Mixmatch: a holistic approach to semi-supervised learning, Adv. Neural Inf. Process Syst. 32 (2019).

[36] S.C. Wong, A. Gatt, V. Stamatescu, M.D. McDonnell, Understanding data augmentation for classification: when to warp?, in: 2016 international conference on digital image computing: techniques and applications (DICTA), IEEE, 2016, pp. 1–6.

[37] Y. Tokozume, Y. Ushiku, T. Harada, "Between-class learning for image classification, in: roceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 5486–5494.

[38] E. Arazo, D. Ortego, P. Albert, N. O'Connor, K. McGuinness, Unsupervised label noise modeling and loss correction, in: International conference on machine learning, PMLR, 2019, pp. 312–321.

[39] I. Goodfellow, et al., Generative adversarial nets, Adv. Neural Inf. Process Syst. 27 (2014).

[40] A. Regev, et al., Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection, Am. J. Gastroenterol. 97 (10) (2002) 2614–2618.

[41] V. Ratziu, et al., Sampling variability of liver biopsy in nonalcoholic fatty liver disease, Gastroenterology 128 (7) (2005) 1898–1906.

[42] C. Caussy, et al., Association between obesity and discordance in fibrosis stage determination by magnetic resonance vs transient elastography in patients with nonalcoholic liver disease, Clin. Gastroenterol. Hepatol. 16 (12) (2018) 1974–1982.

[43] N. Kawamura, et al., Influence of liver stiffness heterogeneity on staging fibrosis in patients with nonalcoholic fatty liver disease, Hepatology 76 (1) (Jul. 2022) 186–195, https://doi.org/10.1002/hep.32302.

[44] R. Heide, D. Strobel, T. Bernatik, R.S. Goertz, Characterization of focal liver lesions (FLL) with acoustic radiation force impulse (ARFI) elastometry, Ultraschall in der Medizin-European J. Ultrasound (2010) 405–409.

[45] C.H. Suh, et al., Shear wave elastography as a quantitative biomarker of clinically significant portal hypertension: a systematic review and meta-analysis, Am. J. Roentgenol. 210 (5) (Mar. 2018) W185–W195, https://doi.org/10.2214/AJR.17.18367.

[46] T. Deffieux, et al., Investigating liver stiffness and viscosity for fibrosis, steatosis and activity staging using shear wave elastography, J. Hepatol. 62 (2) (2015) 317–324, https://doi.org/10.1016/j.jhep.2014.09.020.

[47] J. Bercoff, M. Tanter, M. Fink, Supersonic shear imaging: a new technique for soft tissue elasticity mapping, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 51 (4) (2004) 396–409.

[48] V. de Lédinghen, J. Vergniol, Transient elastography (FibroScan), Gastroenterol. Clin. Biol. 32 (6 SUPPL. 1) (Sep. 2008) 58–67, https://doi.org/10.1016/S0399-8320(08)73994-0.

[49] G. Ferraioli, et al., WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 3: Liver, Ultrasound Med. Biol. 41 (5) (May 2015) 1161–1179, https://doi.org/10.1016/j.ultrasmedbio.2015.03.007.

[50] Y. Deng, N.C. Rouze, M.L. Palmeri, K.R. Nightingale, Ultrasonic Shear Wave Elasticity Imaging (SWEI) Sequencing and Data Processing Using a Verasonics Research Scanner, doi: 10.1109/TUFFC.

[51] T. Shiina, et al., WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 1: basic principles and terminology, Ultrasound Med. Biol. 41 (5) (May 2015) 1126–1147, https://doi.org/10.1016/j.ultrasmedbio.2015.03.009.

[52] M. Friedrich-Rust, et al., Performance of transient elastography for the staging of liver fibrosis: a meta-analysis, Gastroenterology 134 (4) (2008) 960–974.

[53] G. Hinton, O. Vinyals, J. Dean, Distilling the knowledge in a neural network, arXiv preprint arXiv:1503.02531, 2015.